Title

Effect of earthquake-related losses and post-earthquake events on morbidity and mortality: causal mediation analysis of the prospective cohort data of the 1988 earthquake survivors in Armenia

Author

Khachadourian, Vahe

Publication Date

2019

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CLIFORNIA

Los Angeles

Effect of earthquake-related losses and post-earthquake events on morbidity and mortality: causal mediation analysis of the prospective cohort data of the 1988 earthquake survivors in Armenia

A dissertation submitted in partial satisfaction of the

requirements for the degree Doctor of Philosophy

in Epidemiology

by

Vahe Khachadourian

2019

ABSTRACT OF THE DISSERTATION

Effect of earthquake-related losses and post-earthquake events on morbidity and mortality: causal mediation analysis of the prospective cohort data of the 1988 earthquake survivors in Armenia

by

Vahe Khachadourian

Doctor of Philosophy in Epidemiology

University of California, Los Angeles, 2019

Professor Onyebuchi Aniweta Arah, Co-Chair

Professor Haroutune K. Armenian, Co-Chair

Earthquakes are one of the most common types of natural disasters. They remain unpredictable and often result in substantial damage and destruction. Earthquakes have been found to be associated with adverse mental and psychological conditions such as post-traumatic stress disorder, depression, and generalized anxiety disorder. Despite the significant negative health and social impacts of earthquakes, to date, only a limited number of studies have explored physical health outcomes and long-term survival among populations exposed to earthquakes. Most of these studies have methodological limitations such as lack of generalizability due to population specificity; limited sample size and power; non-temporal design; and the use of aggregate-level exposures and outcomes. In this dissertation, we presented a few causal mediation analysis frameworks and discussed their potential in exploring causal mechanisms in

disaster research. Subsequently, we used data from a prospective cohort of surviving adults with differential exposure levels to the 1988 earthquake in Armenia to investigate the impact of earthquake-related exposure (housing damage, death of a family member, and serious injury) on all-cause mortality. Various data adaptive approaches, including super learner and random survival forests algorithms were applied to simulate and impute the outcome for the subpopulation with a shorter follow-up time. Furthermore, we applied the four-way mediation analysis framework and decomposed the effect of housing damage on all-cause mortality with respect to receiving permanent housing in the aftermath of the earthquake. Similarly, we assessed the impact of earthquake-related exposure on risk of developing diabetes. The interventional approach for path-specific effect estimation framework was used to decompose the effect of housing damage on diabetes with respect to permanent housing aid and job loss. Finally, a simulation study, incorporated the estimated effects to assess the impact of hypothetical interventions on risk of all-cause mortality and diabetes. The hypothetical interventions were based on providing permanent housing, or preventing job loss in the aftermath of the event. The dissertation sheds light on effects of earthquake related exposure and potential post-earthquake interventions on risk of all-cause mortality and diabetes. Furthermore, it provides accessible evidence for decision making and prioritizing interventions and optimal resource allocation in the aftermath of an earthquake. Future studies can expand the current work by investigating the effect of earthquake related exposure through other pathways. Such studies would also benefit from the inclusion of social support and psychological outcomes in their framework analysis. Future studies should also consider more extensive sensitivity analysis for uncontrolled confounding and incorporate sensitivity analysis for measurement error and selection bias.

The dissertation of Vahe Khachadourian is approved.

Ronald S. Brookmeyer

Karin B. Michels

Onyebuchi Aniweta Arah, Committee Co-Chair

Haroutune K. Armenian, Committee Co-Chair

University of California, Los Angeles

2019

Dedication

To my family

Table of Contents

List of Figures

List of Tables

List of Abbreviations

AUROC	Area under the receiver operating characteristic
BMI	Body mass index
CDE	Controlled direct effect
C-index	Concordance index
DSM-III	Diagnostic and Statistical Manual of Mental Disorder
ID	Unique identifier
MIE	Mediated interaction effect
NDE	Natural direct effect
NIE	Natural indirect effect
PAI	Portion attributable to interaction
PDE	Pure direct effect,
PEPSI	Post-earthquake psychopathological investigation
PIE	Pure indirect effect
PTSD	post-traumatic stress disorder
RIE	Reference interaction effect
SEM	Structural equation modeling
SUTVA	Stable unit treatment value assumption
TDE	Total direct effect
TE	Total effect
TIE	Total indirect effect

Acknowledgments

My academic endeavor and PhD studies at UCLA mark an important milestone in my professional and personal development. This experience would not have been possible, let alone fruitful, if I did not have the unconditional support of my family, friends, colleagues, mentors, and faculty of the Department of Epidemiology at the Fielding School of Public Health.

I am deeply indebted to Dr. Haroutune Armenian, my mentor and co-chair of my dissertation committee, for his dedication and unconditional support. As the professor of my first epidemiology course (back in 2009), you introduced me to the field. Since then, I have continued to benefit from your wisdom, encouragement, support, and perspective. Your wealth of expertise inspires me and fuels my passion for the subject. You are an outstanding epidemiologist and a wonderful human being, and I am very fortunate to have had the opportunity to work with and learn from you.

I would also like to express my deepest gratitude to my mentor and co-chair of my dissertation committee, Dr. Onyebuchi Arah for your guidance, inspiration, and encouragement throughout my journey at UCLA. I cannot thank you enough for your generosity with your time and knowledge, and for introducing me to the world of causal inference and quantitative research methodology. I am also very grateful for the many educational and scientifically-simulating discussions we had. Your teaching style, dedication to student learning, and enthusiasm are exemplary and have been empowering. You will always be a role model to me.

I am also grateful to my committee members, Dr. Brookmeyer and Dr. Michels. Thank you so much for agreeing to be on my committee and providing your invaluable time and feedback during my dissertation studies.

My PhD studies and adventure at UCLA would not have been possible without the sacrifice, advice, support, encouragement, and love of my parents, Vigen and Knarik, and my sister Melani. I also want to thank Yeraz and my friends for cheering me on and standing by my side throughout my journey. Yeraz, thank you for your unconditional love and support.

I greatly appreciate and acknowledge the generous funding support for my PhD studies from the Turpanjian Family Educational Foundation, the Calouste Gulbenkian Foundation, the Department of Epidemiology at UCLA, and the UCLA Graduate Division.

Chapter 2 is partly incorporated from Khachadourian V, Armenian HK, Demirchyan A, Melkonian A, Hovanesian A. Post-Earthquake Psychopathological Investigation in Armenia: research methodology, summary of previous findings and recent follow-up. *Disasters.* 2016, Jul. 40(3):518-33. Doi: 10.1111/disa.12166.

Vita/Biographical Sketch

Education

2004 – 2011	Doctor of Medicine (MD) Yerevan State Medical University, Yerevan, Armenia
2009 – 2011	Masters of Public Health (MPH) American University of Armenia, Yerevan, Armenia

Publications

Goenjian AK, **Khachadourian V**, Armenian HK, Demirchyan A, Steinberg AM. Posttraumatic Stress Disorder 23 Years after the 1988 Spitak Earthquake in Armenia. *Journal of Traumatic Stress*. 2017.

Grigoryan R, Petrosyan V, Melkomian DM, **Khachadourian V**, McCartor A, Crape B. Risk factors for children's blood lead levels in metal mining and smelting communities in Armenia: a cross-sectional study. *BMC Public Health.* 2016, Sep. 16:945.

Movsisyan A, Demirchyan A, **Khachadourian V**, Armenian HK, Diener-West M, Goenjian A. Diagnostic Accuracy and Operating Characteristics of the Posttraumatic Stress Disorder (PTSD) Checklist in the Postearthquake Population in Armenia. *Journal of Traumatic Stress Disorders & Treatment*. 2016, 5:1.

Khachadourian V, Armenian HK, Demirchyan A, Melkonian A, Hovanesian A. Post Earthquake Psychopathological Investigation in Armenia: research methodology, summary of previous findings and recent follow-up. *Disasters.* 2016, Jul. 40(3):518-33.

Demirchyan A, Goenjian A, **Khachadourian V**. Factor Structure and Psychometric Properties of the Posttraumatic Stress Disorder (PTSD) Checklist and DSM-5 PTSD Symptom Set in a Long-Term Post earthquake Cohort in Armenia. *Assessment.* 2015, Oct. 22(5):594-606.

Demirchyan A, Petrosyan V, Armenian HK, **Khachadourian V**. Prospective study of predictors of poor self-rated health in a 23-year cohort of earthquake survivors in Armenia**.** *Journal of Epidemiology and Global Health.* 2015, Sep. 5(3):265-74.

Khachadourian V, Truzyan N, Harutyunyan A, Thompson ME, Harutyunyan T, Petrosyan V. People-centered tuberculosis care versus standard directly observed therapy: study protocol for a cluster randomized controlled trial. *Trials*. 2015, Jun. 16:281.

Khachadourian V, Armenian HK, Demirchyan A, Goenjian A. Loss and Psycho-social Factors as Determinants of Quality of Life in a Cohort of Earthquake Survivors*. Health and Quality of Life Outcomes. Health and Quality of Life Outcomes*. 2015, Feb. 6;13:13.

Demirchyan A, **Khachadourian V**, Armenian HK, Petrosyan V. Short and Long Term Determinants of Incident Multimorbidity in a Cohort of 1988 Earthquake Survivors in Armenia. *International Journal for Equity in Health*. 2013, Aug.12(1):68.

Chapter 1. Introduction

1.1 Epidemiology of disasters

Natural disasters are unpredictable and generally non-preventable. According to a recent report by the United Nations, during the period of 1994–2013, natural disasters have affected 4.4 billion people and have killed more than 1.3 million. The economic cost of disasters during the same period has surpassed 2 trillion USD (Kellet, 2014). While specific outcomes vary, the health, environmental, and economic consequences are invariably severe. Table 1 summarizes losses associated with natural disasters between 2009 and 2013.

Table 1.1 Losses associated with natural disaster during the period of 2009-2013 by year

Year	# of Disasters	# of People Killed	# of People Affected	Economic Damages, in $USD
2009	335	10,655	> 119 million	> $41 billion
2010	385	297,000	> 217 million	> $123 billion
2011	332	30,773	> 244 million	> $366 billion
2012	357	9,655	> 124 million	> $157 billion
2013	330	21,610	> 96 million	> $118 billion

Disasters have disproportionately greater impacts in low- and lower-middle income countries. For instance, while only 33% of the countries impacted by natural disasters were low- or lower-middle income, they suffered more than 80% of the disaster-related deaths (Kellet, 2014). Still, middle- and high-income countries also experience severe effects. (Binder & Sanderson, 1987) The Tohoku Earthquake resulted in severe damage in Japan, a high-income country with a very high level of disaster preparedness. According to official statistics from the Japan National

Police Agency, the disaster resulted in more than 15,500 deaths and more than 2,500 people missing (as of June 10, 2016). The total cost of the disaster was estimated to be $90–250 billion USD (Kazama & Noda, 2012).

1.2 Health consequences of disasters

Natural disasters are a major public health concern, and the short-term negative impacts of disasters on physical and mental health outcomes are well established in the scientific literature. Post-disaster mental-health research has a long and interesting history (Andreasen, 2010). Researchers have described various psychopathologies among populations exposed to traumatic events (Adler, 1943; Grinker, 1944; Kral, 1951). The third edition of the Diagnostic and Statistical Manual of Mental Disorders (DSM-III) was the first reference text to formalize post-traumatic stress disorder (PTSD) as a pathology (American Psychiatric Association, 1980). Since the publication of DSM-III in 1980, many studies have examined PTSD rates and determinants among populations affected by disasters. PTSD is one of the most frequently studied and perhaps the most commonly occurring psychopathology in the aftermath of disasters (Armenian et al., 2000; Foa, Stein, & McFarlane, 2006; Goldmann & Galea, 2014; Norris et al., 2002). Most of the studies targeting populations of natural disasters have documented PTSD prevalence from as low as 5% to as high as 70%. A study of 573 adult survivors living in shelters after the 1985 Mexico City earthquake reported a PTSD prevalence of 32% within 10 weeks after the event (de la Fuente, 1990). Similarly, a study of survivors living in prefabricated housing after the 1999 earthquake in Marmara found a PTSD prevalence of 39% within a mean of 20 months after the event (Salcioglu, Basoglu, & Livanou, 2003). In the aftermath of a devastating tornado in North Carolina in 1984, Madakasira and O'Brien (1987) found that about 59% of adults surveyed met the PTSD diagnostic criteria. Various studies have also shown that severity of exposure in a disaster is a significant predictor of psychopathology in

the aftermath of a traumatic event. For instance, in a study of adult populations exposed to the 1988 Spitak earthquake in Armenia, Armenian et al.(2000) found that those with higher levels of financial and material loss had higher PTSD diagnosis rates.

Although studies indicate that the rates of PTSD and other post-disaster psychopathologies decline with time, they typically find that, years later, the prevalence rates of psychopathologies are higher than in comparison-study populations of individuals who did not experience the disaster (A K Goenjian et al., 2000; McFarlane, 1988). A study by Hull et al. (2002) reported that PTSD among survivors of the Piper Alpha oil fire disaster declined from 73% after three months to 21% 10 years after the disaster. Despite this decline, the observed rates at 10 years were still higher than those in the general population. Similarly, in a follow-up study conducted among survivors of the 1988 Spitak earthquake survivors in Armenia, the PTSD rates attributable to the earthquake decreased from 48% in 1991 to less than 12% in 2012 and remained a significant health problem among those living in areas severely impacted by the earthquake (Armen K. Goenjian, Khachadourian, Armenian, Demirchyan, & Steinberg, 2018). Morgan et al. (2003) found that 33 years after the Aberfan disaster in South Wales, where a coal slag collapsed on a primary school, 29% of survivors still had current PTSD. This study also found that the disaster survivors had more than three times the odds of developing PTSD than the control group.

In contrast to these studies, in a follow-up study 8 years after the 1997 Marche earthquake in Italy, Priebe et al. (2011) found no long-term negative mental-health impact among survivors of the earthquake. After 17 years of follow-up, Green et al. (1990) found that among survivors of the Buffalo dam collapse disaster, survivors that were children during the disaster had significantly higher rates of lifetime PTSD, but their rates of current PTSD, major depression, and general anxiety disorders were not significantly different from the comparison group. Some of the variation in rates and trajectories of these psychopathologies within the literature may be

due to differences in the nature and intensity of traumatic events, sociocultural and demographic differences in the study populations, and methodological differences and limitations in study designs and conduct (Neria, Galea, & Norris, 2009; Norris, 2006).

PTSD is not the only post-disaster psychopathology that has generated intense interest from researchers. The risk for several other psychopathologies has been found to increase in the aftermath of disasters (Armenian et al., 2002; Goldmann & Galea, 2014; Guo, He, Qu, Wang, & Liu, 2017; Khachadourian, Armenian, Demirchyan, Melkonian, & Hovanesian, 2016; Neria et al., 2009).

1.3 Gaps in the literature

Studies suggest that the intensity of a victim's exposure in a disaster is directly correlated with the risk or severity of negative health outcomes. Armenian et al. (2000) was the first to establish a dose-response relationship between the amount of loss to a family and PTSD. They also observed a similar predictive pattern for the development of depression after a disaster (Armenian et al., 2002). The disaster-related risk factors for PTSD have also found to be responsible for other mental-health conditions among disaster survivors (e.g. depression and general anxiety disorder) (Goldmann & Galea, 2014; Neria et al., 2009). Although DSM categorizes these pathologies as separate disorders, they have similar manifestations and share common symptoms; they also have high comorbidity rates, rarely occurring in isolation (Brady, Killeen, Brewerton, & Lucerini, 2000). The high rates of co-occurrence, the common symptoms, and shared diagnostic criteria of these mental health conditions suggest variations across a single pathological spectrum, rather than separate groups of diseases (Watson, 2005). The regular — and often notable — changes introduced to the diagnostic criteria of PTSD since its introduction in 1980 are another indication that our understanding of post-trauma reactions is

evolving rapidly and may be incomplete. The most recent changes in PTSD criteria have sparked a heated debate among the scientists focusing on the issue (Demirchyan, Goenjian, & Khachadourian, 2015; Pai, Suris, & North, 2017).

Several studies have examined the association between PTSD and physical health outcomes. Despite some variations in methodology and result of studies, several prospective studies have demonstrated that PTSD is positively associated with increased risk of cardiovascular diseases (Beristianos, Yaffe, Cohen, & Byers, 2016; Boscarino, 2006; Gradus et al., 2015; Jordan, Miller-Archie, Cone, Morabia, & Stellman, 2011) and diabetes (Boyko et al., 2010; Roberts et al., 2015). Nevertheless, in light of potential methodological challenges and limitations, including uncontrolled confounding, selection bias, and reverse causality (K. C. Koenen et al., 2017), observed associations of PTSD with cardiovascular diseases and diabetes have often not been inferred as causal associations. These circumstances have prevented establishing PTSD as a risk factor for cardiovascular diseases and diabetes (K. C. Koenen et al., 2017).

The direct impacts of disaster exposure on physical health and mortality are less well studied, and longer-term studies are especially scant. Circulatory and nervous-system diseases are the most common focus of studies assessing physical health in the aftermath of earthquakes (Armenian, Melkonian, & Hovanesian, 1998; Hata, Akiyama, Wakui, Takasaka, & Shiono, 2012; Kamoi, Tanaka, Ikarashi, & Miyakoshi, 2006; Kario, Matsuo, Shimada, & Pickering, 2001; Kario & Ohashi, 1997; Nakano et al., 2012; Ripoll Gallardo et al., 2018; Sofia et al., 2012; Sokejima et al., 2004; Tsuchida et al., 2009; Wu, Cheung, Cole, & Fink, 2014).

Disaster studies face difficult methodological limitations because of the inherent logistical challenges present in a disaster-stricken area. These include lack of generalizability due to population specificity; limited sample size and power; non-temporal design; and the use of

aggregate-level exposures and outcomes. Moreover, the majority of the few studies available that focus on physical health outcomes have been conducted in high-income countries; despite the higher vulnerability of low- and middle-income countries, the physical-health impacts in those countries have received little attention.

Earthquakes constitute a large proportion of the death toll and economic loss caused by natural disasters. In 2015, more than 30% of the 22,700 natural disaster related deaths were attributable to earthquakes. The trend of continual population growth and migration to urban areas suggests that the impact of earthquakes on human populations will only increase. The 2010 earthquake in Haiti resulted in extensive destruction, with various sources estimating deaths from 222,000 to 316,000. The Haiti quake injured 300,000 and destroyed 97,000 houses, displacing 1.3 million from their areas of residence (U.S. Geological Survey, 2014).

Considering the limited research on long-term effects of natural disasters and earthquakes in particular on physical health and mortality, this dissertation proposes a conceptual framework and aims to assess the independent impact of selected earthquake exposures on physical health and mortality. Knowledge about the long-term effects of earthquake exposure and post-earthquake events on health outcomes can inform future policies, allowing for more effective planning and response.

1.4 Conceptual framework

Reviewing the literature on disaster epidemiology, we developed and conceptualized a general framework linking impacts of disaster and post-disaster experiences on health outcomes in the context to pre-, peri-, and post-disaster factors (Figure 1.1). Although the role of pre-disaster risk factors such as gender, socioeconomic status, education, and previous health are well

documented, the current framework focuses on investigating the impact of selected factors from peri- and post-disaster phases of the process. We hypothesized that exposure to natural disaster has direct and indirect negative consequences on health outcomes, where the post-disaster factors act as the main mediators (or part of mechanisms) for the indirect effect. Figure 1.1 presents the proposed conceptual framework.

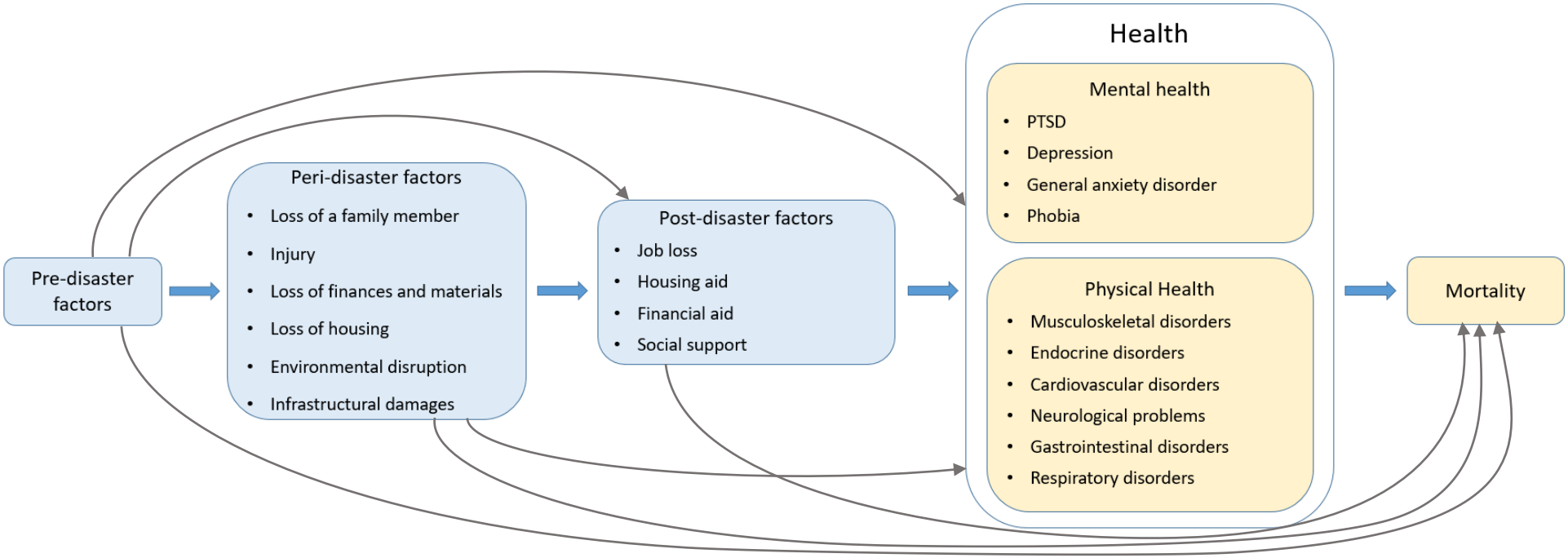

Figure 1.1 Conceptual framework for the effects of natural disasters on health outcomes.

1.5 Mediation analysis and its importance in disaster research

An indispensable statistical framework used across many disciplines, mediation analysis examines possible relationships between two correlated variables by positing a third, intervening variable. This hypothetical “mediator” allows researchers to explore beyond direct causal links when investigating potential effect relationships (Hafeman & Schwartz, 2009; Kenny, 2008). Mediation analysis can be used to test mediating pathways and mechanisms in both experimental and observational research (MacKinnon, Kisbu-Sakarya, & Gottschall, 2013).

The concept of mediation also provides a ground for causal inference. For instance, identification of a hypothesized mediator can provide supporting evidence for the existence of a causal relation between the observed exposure and outcome (Hafeman & Schwartz, 2009). Mediation analysis also enables researchers to distinguish true null effects from non-null, pathway-specific effects averaging to a total null effect (MacKinnon, 2008). In extreme cases, one or a few pathway-specific effects can be at odds with the total effect of exposure. Identification of such scenarios and pathways can help not only in better understanding the risk factors of a disease, but also in developing effective interventions.

A primary objective of public-health research is to supply the evidence for policy decision-making that is designed to improve health outcomes. However, the knowledge produced by science often cannot be readily utilized in policy-making (Judea Pearl, 2014b). Mediation analysis is a framework that can offer evidence and more accessible information for policy-making. The evidence provided by mediation analysis is even more robust when simulations are used; these can provide quantitative estimates for effect of potential interventions on outcome. Mediation can also guide the refinement of existing interventions by targeting specific pathways (Vanderweele, 2015). In some circumstances, even when it might not be feasible or ethical to

directly intervene in the exposure, knowledge about mediating mechanisms of the exposure-outcome effect can provide alternative options to target the effect of such exposures.

Disasters — particularly natural disasters — are often sudden and unexpected. Therefore, minimizing their immediate impact requires advanced long-term planning, and significant financial and scientific investment in infrastructure and disaster preparedness. Lack of such developed infrastructure in low-income, middle-income, and even many high-income countries makes populations more vulnerable to adverse effects of disasters. Even a well-prepared country can easily experience disaster impacts that exceed their level of preparedness.

Since affected communities inevitably experience some level of negative impact, it's crucial that researchers and policy-makers gain a better understanding about how post-disaster factors can mediate the impact of long-term health outcomes. Such knowledge can help policy-makers develop targeted strategies to mitigate negative consequences for disaster-affected communities, particularly for those most impacted.

Hence, this dissertation aims to conduct a mediation analysis to investigate the role of post-earthquake factors as potential mediators.

1.6 Structure and specific aims of this dissertation

This dissertation will investigate the impact of earthquake-related losses and post-earthquake events as mediators on morbidity and mortality for a cohort of 1988 earthquake survivors in Armenia using data from the **p**ost-**e**arthquake **ps**ychopathological **i**nvestigation (PEPSI).

The objectives and specific aims of each chapter are summarized below.

Chapter 2: Material and methods

This chapter presents the cohort population and the methods used to answer the research questions. It continues with description of the methods used to prepare the dataset for analysis. We discuss the classical approach to causal mediation analysis and review the more recent techniques and developments in the field — particularly those relevant to time-to-event settings. The chapter concludes by identifying and providing justification for the method and framework of mediation analysis to be used in the subsequent chapters.

Chapter 3: Impact of earthquake exposure on and the mediating role of post-earthquake experiences in all-cause mortality

In this chapter, we investigate the effect of earthquake exposure and loss on all-cause mortality. Further, we decompose the total effect of earthquake-related loss on all-cause mortality with respect to post-earthquake support. The specific research questions are:

a) What are the effects of earthquake-related losses — such as the death of a nuclear family member, injury, and loss of housing — on all-cause mortality?

b) What is the potential mediating or mitigating role of providing permanent housing in the potential effect of loss of housing on all-cause mortality?

Chapter 4: Impact of earthquake exposure on and the mediating role of post-earthquake experiences in incident diabetes

Chapter 4 investigates the effect of earthquake exposure and loss on diabetes. Similarly, it explores the potential effect of earthquake-related loss by decomposing it into direct and indirect

effects with respect to providing new permanent housing and job loss in the aftermath of the disaster.

Chapter 5: Comparison of hypothetical interventions to reduce post-earthquake morbidity and mortality

Chapter 5 evaluates and quantifies the potential effect of realistic hypothetical interventions [targeting the identified factors and mechanisms in Chapters 3 and 4] on morbidity and mortality among the earthquake survivors using parameters and effect estimates supplied from Chapters 3 and 4.

Chapter 2. Materials and Methods

2.1 Methods and protocol of PEPSI study (source of data)

Study setting, summary of the earthquake, and PEPSI

PEPSI was designed to probe the short-and long-term mental and physical health effects of the 6.9 magnitude (on the Richter scale) earthquake that struck the northern part of Armenia on 7 December 1988, damaging almost one-third of the country. The closest city to the epicenter of the earthquake, Spitak, was totally destroyed. The earthquake was responsible for the deaths of more than 25,000 people, the injuries of some 100,000 people, and the loss of housing of approximately 500,000 people (Hadjian, 1993). The overall cost of the damage was estimated at more than USD 16 billion. The damage to buildings and the electricity and water supply infrastructure, along with harsh climatic conditions, contributed to a further deterioration in the living circumstances of survivors.

The PEPSI, conducted between 1990 and 2012, was a study of a prospective cohort of the survivors of the Spitak earthquake. It aimed to probe the trajectories and determinants of potential short- and long -term physical and mental health consequences, their impact on health-related quality of life, and effective resilience mechanisms to overcome the adverse effects of the disaster. Four phases of the PEPSI have been completed to date.

Study population

The study cohort was composed of employees of the Ministry of Health of the Republic of Armenia in the Soviet Union and their first-degree relatives who were living in the earthquake region on the day prior to the earthquake on 7 December 1988. The method for selecting the cohort yielded a large and diverse study population, with employees' personnel files providing a

means to follow-up and track participants. As noted, the study has involved four phases to date—there is a possibility of additional follow-up phases in the future. During Phase I, a baseline stage that started in April 1990 and ended in December 1992, 7,016 of the 9,017 employees of the Ministry of Health at the time of the earthquake were interviewed. With the inclusion of family members, the size of the total study cohort rose to 32,743 individuals (see Armenian et al. (1997) for further details on the baseline sampling method).

During data collection for Phase I (from early 1991 until later in the same year), Phase II psychological assessments were performed for a geographically stratified sub-sample of 1,785 individuals with a minimum age of 18 years (Armenian et al., 2000, 2002). This sub-sample was stratified to over-sample survivors from areas with extensive earthquake damage; however, there were no significant differences in the initial demographic and the basic characteristics of this sub-sample and that of the same age group of the baseline total cohort (Khachadourian et al., 2016). Phase III, a follow-up evaluation of Phase I participants, occurred approximately two years after the first phase for each contributor, during 1992–94 (Armenian et al., 1998). Phase IV started in April 2012 (almost 21 years after the baseline study) and lasted for two months, involving follow-up of the stratified sub-sample of participants who underwent a psychological assessment in Phase II and who were also followed successfully during Phase III. Phase IV was launched after a small-scale pilot project (Khachadourian, 2011) revealed that it was feasible to follow-up on this population after two decades.

Figure 2.1 presents the flow chart of population availability and follow-up during the different phases of this study. From the initial sample of 9,017 employees of the Ministry of Health, 7,016 were located and consented to participate in Phase I. Because participants' families were also included in the study, the total study population consisted of 32,743 individuals. Phase II

approached a sub-sample of 1,785 individuals. During the Phase III follow-up, 97% of the original cohort including 99% of Phase II sub-sample participated (1,773 individuals).

Contact information to locate the Phase IV participants was obtained from the baseline database and the national election registry (a publicly accessible database of citizens). At Phase IV, 23-years after the initial phase, 1487 (84%) of the original sub-sample of 1773 (excluding 12 individuals who were lost to follow-up during Phase III) were located. Among the Phase II sub-sample more than 80% of intended participants or their families (if the intended participant had died, moved out of the country, or was unable to participate because of severe health conditions) consented to contribute. Out of those located, 309 (21%) were dead, 300 (20%) had permanently moved out of country, 89 (6%) were not able to participate (mainly because of severe health conditions or temporarily not being available), 725 (49%) assented to participate and 64 (4%) refused. The study was able to obtain some survival status and if applicable mortality dates for 1487 (83%) of the initial 1773 sample.

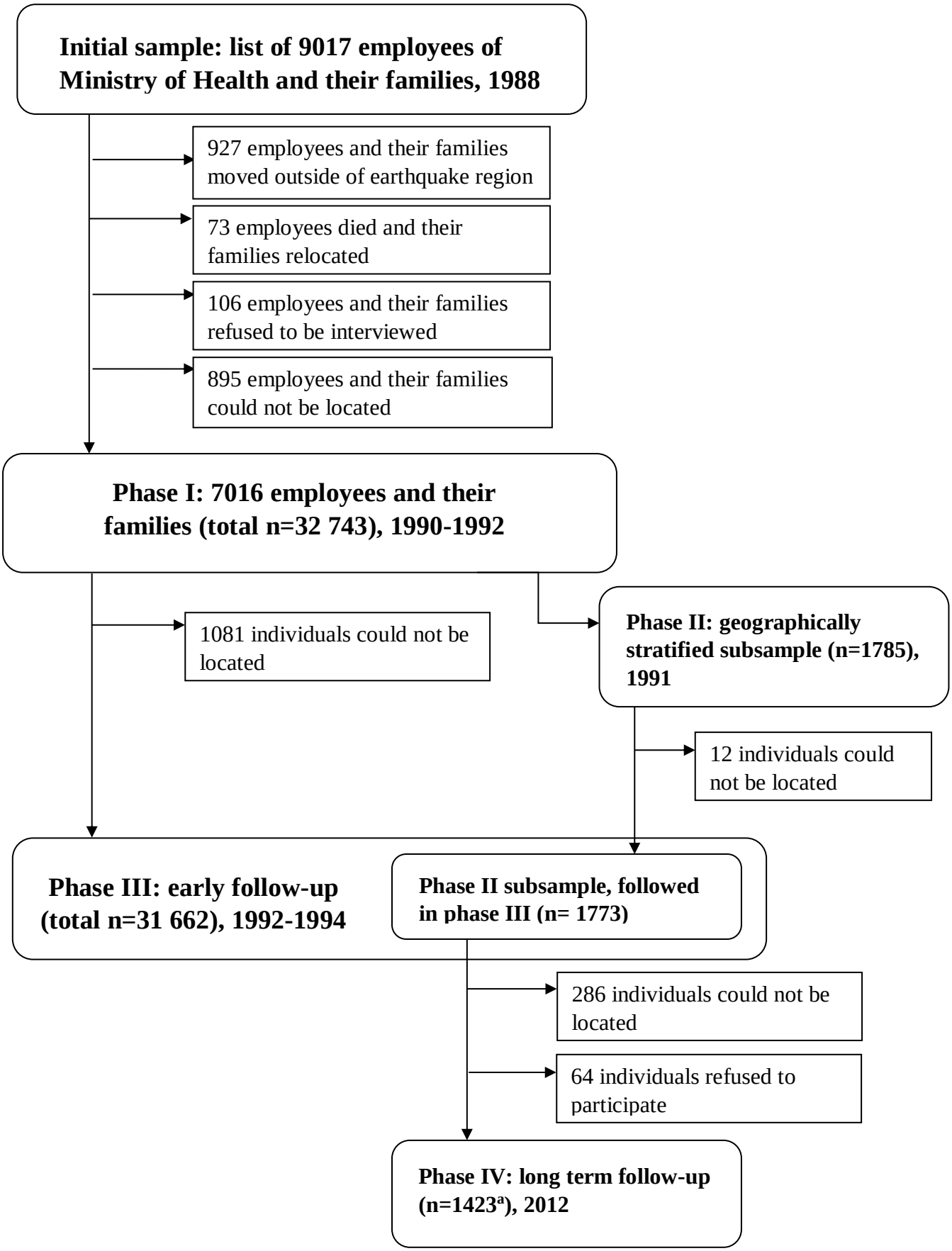

Figure 2.1 Sample of survivors initially and throughout follow-up during Phases I–IV

[a] Among these 1423, there were 722 main participants who answered the questionnaire. 330 individuals had died and 348 individuals were either out of country or unable to participate, for whom we have a proxy questionnaire filled out.

Measurements[1]

A wide range of data were collected during Phase I of the study on losses and damage caused by the Spitak earthquake, changes in employment circumstances after the event, demographic characteristics, family structure, health behaviors, help and support received, standard of living at the time of the earthquake, mortality in the family before, during, and after the earthquake, self-reported height, weight, and physical health, and socioeconomic status. Phase II focused on assessing the psychological condition of the survivors, using collected information on depression, general anxiety, panic symptomatology, phobia, and PTSD—all of these instruments were developed based on the third edition of the Diagnostic and Statistical Manual of Mental Disorders (American Psychiatric Association, 1987), which was current at the time. In addition, questions were posed on social support and interactions (see Armenian et al. (2000 and 2002) for more details about the origin and the development of these instruments). Phase III involved the compilation of self-reported information on mortality and causes of deaths in the family, as well as changes in employment status, health behaviors, and height, weight, and physical health outcomes.

In Phase IV, the study team developed two questionnaires. The first was the main questionnaire designed for use with earthquake survivors and composed of interviewer and self-administered sections. The second was an interviewer-administered proxy questionnaire to collect data on those individuals who died in the interim, were out of the country, or were unable to participate because of severe health conditions. The main questionnaire gathered data on respondent's demographic characteristics, employment history, family structure, health behaviors (such as

[1] Presentation and description of several measurements not relevant for this thesis has been omitted. For further detail on cohort description and methods see: Khachadourian, V., Armenian, H., Demirchyan, A., Melkonian, A. and Hovanesian, A. (2016), A post-earthquake psychopathological investigation in Armenia: methodology, summary of findings, and follow-up. Disasters, 40: 518–533

alcohol consumption, dietary habits, drug abuse, exercise, and smoking), health care utilization, living place history, main and traumatic life events (including dates and their impacts), and socioeconomic status. The study also compiled data on post-earthquake mortality and causes of death in the family.

The proxy questionnaire procured data on unavailable participants' alcohol consumption, demographic characteristics, employment history, self and family members' physical health, smoking pattern, and socioeconomic status, as well as causes of death among family members. As appropriate, data were also collected on the intended participant's cause and year of death, or year and place of migration, or other reason for not being able to participate.

2.2 Methods used in this dissertation

Study population and variables

We used data from the PEPSI to answer the research questions. The analyses were limited to individuals who were at least 18 years old at the time of the earthquake and either had been followed for at least 3 years or had the event (death). The exclusion of children was to avoid mixing potential differential effects of earthquake on health outcomes among children and adults.

Measurements

Data on age, gender, education, employment, pre-earthquake residence location, type of residence (urban vs. rural), and living standard were collected at Phase I and were verified during the subsequent follow-up phases. Age was a continuous variable recorded in years. Gender was a categorical variable. Education was recorded as a continuous variable reflecting the number of years of formal education. The pre-earthquake employment variables included a

binary variable indicating whether the participant was employed or not and a categorical variable pertaining job of those employed. The relative living standard was a self-reported measure with below average, average, and above average categories. Data on pre earthquake life style measures including consumption of alcohol, smoking habits, and physical activity were also systematically collected at various phases of the study. Participants also reported their weight and height before the earthquake. Self-reported health profile of participants at the time of the earthquake encompassed data on cardiovascular diseases including hypertension, diabetes mellitus, stroke, asthma, chronic respiratory diseases, gastrointestinal tract diseases, urinary tract and kidney problems, arthritis, migraine, allergy, cancer, and mental health problems.

Details on earthquake-related injury collected at Phase I of the study included the body part affected and type of injury. We created a binary variable (yes/no) to summarize earthquake-related serious injury to the study participant. Scratches and minor cuts were not considered as a serious injury. Number of nuclear family deaths due to earthquake was collected and since there were only a few individuals with multiple deaths in the family, the variable was dichotomized to no death vs. any death. Housing damage variable was a categorical variable with no damage, minor damages/non-structural damages, and extensive damage/structural damages/total destruction. Data on job loss and remaining unemployed for more than 3 months and receiving a permanent housing during the 3-year period after the earthquake (1989-1991) were also recorded.

At each phase, data on survival outcome of the study participants were collected and for deceased subjects' time of death in years were recorded. At Phase IV, when a first degree relative was available, a proxy questionnaire was administered to collect data on survival outcome of intended study participants who were out of country, unable to participate, or

deceased at the time of Phase IV study. The proxy questionnaire captured data on the outcome for the time period even after the intended participant left the country. In case of proxies not aware of the intended participant's survival status the participants were considered as censored at the time of migration.

Data preparation

The dataset included a total of 26,478 adult participants. To assure full availability of data on receiving a permanent housing as well as job loss during the 3 years following the earthquake (1989-1991), all participants with a censoring time before the end of 1991 were excluded from the analytical sample, and the final analytical sample totaled to 23,639.

As described above, the majority of the sample was followed up to Phase III (approximately 4 years after the earthquake), and longer-term follow-up information was only obtained for a sub-sample of individuals who were successfully followed at Phase IV. To maximize the use of information from observations administratively censored at phase I or III of the study or non-administratively censored (lost to follow-up) at any stage of the follow-up, we applied a machine learning (data adaptive) approach, the super learner, a type of ensemble learner to obtain the survival outcome for the participants during the period that were not observed. As sensitivity analyses, we additionally applied a) the random survival forest algorithm, an ensemble algorithm, and b) a modern score based multiple imputation approach to predict and impute the survival outcome for those observations censored before the end of the follow-up. Further details about the methods and the description of their application are provided below.

Super learner

Given the existence of potentially multiple generalizers for each learning set, most of prediction algorithms, including cross-validation (Stone, 1977), bootstrapping (Efron, 1979), and generalized cross-validation (Li, 1985), select a single best algorithm from the initial pool of presented algorithms (Wolpert, 1992). Stacked generalization, as proposed by Wolpert (1992), is an ensemble method that addresses this limitation and offers a single prediction algorithm by allowing to combine multiple prediction algorithms (learners). Such a hybrid algorithm can achieve a predictive accuracy exceeding the highest predictive accuracy reached by the individual algorithms included in the ensemble (Wolpert, 1992). Since then, several methods have incorporated the approach, and the approach is commonly referred to as stacking in the literature (Naimi & Balzer, 2018). Van der Laan and colleagues showed that, asymptotically, the stacking approach is guaranteed to perform as well as the best individual algorithm included in the ensemble (Van der Laan & Dudoit, 2003; Van der Laan, Dudoit, & Van der Vaart, 2006; Van der Laan, Polley, & Hubbard, 2007a). The proposed approach benefits from larger number of algorithms included in the ensemble as it creates a weighted combination of the candidate algorithms to benefit from advantages of the best performing algorithms, hence termed as super learner. The algorithm's optimality can be defined as minimization of a specific type of loss function that can be specific to the type of the outcome. Examples of loss functions include the squared error loss function, the rank loss function (aiming to minimize the area under the receiver operating characteristic (AUROC) curve), and the negative log loss function. The V-fold cross-validation method in the super learner helps with preventing over fitting of the data.

The super learner includes one or more algorithms to model the outcome. It divides the entire dataset into *V* mutually exclusive and exhaustive samples with a sample size approximately equal to *N*/*V*, where *N* is the total number of observations in the entire dataset. For instance, performing a *V*-fold cross-validation with $V = 10$ in a super learning procedure will require to divide the study sample into 10 approximately equal sub-samples, v_1 through v_{10}. At the first

iteration, the super learner excludes the first sub-sample, v_1, and applies the first algorithm to the pool of v_2-v_{10} sub-samples (training sets). Then the algorithm is applied to the sample excluded from the analysis, v_1, to obtain the prediction and evaluate the performance of the predictor obtained by applying the algorithm to the pool of v_2-v_{10} sub-samples. The process is repeated 9 more times, each time using one of the sub-samples, v_2-v_{10} sub-samples, as the validation set. The same steps are applied to every other algorithm included in the ensemble. The estimated risk for each algorithm is obtained by averaging the calculated risk from the 10 validation sets. The super learner's function is to obtain an estimator that minimizes the cross-validated risk of candidate algorithms in the ensemble library. Super learner achieves this by calculating an optimal vector of weights and applying them to the candidate algorithms in the ensemble library. Further description of the method (Van der Laan & Dudoit, 2003; Van der Laan et al., 2006), its demonstration (Naimi & Balzer, 2018; Eric C Polley & van der Laan, 2010) and more advanced examples (Luque-Fernandez et al., 2018; Pirracchio et al., 2015; Rose, 2013; Van der Laan, Polley, & Hubbard, 2007b; Zheng, Balzer, van der Laan, & Petersen, 2018) are available.

The off-the-shelf statistical software for applying the method and its related prediction algorithms are mostly for continuous or binary outcomes. To implement the super learner using the available R package (SuperLearner), we restructured the person-oriented time-to-event dataset into a person-period (discrete person time) dataset where each person had one line of observation in the dataset for each discrete time (from the start of the observation up to the event/censoring time). Singer and Willett (1993) provide a detailed discussion of the restructuring of the dataset, code for performing the dataset restructuring, and guide for application of discrete time survival analysis. To allow for flexibility of time-dependent intercepts in the models beyond the linear function provided by the ordinal time since the earthquake [in

years], we created a natural cubic spline with five knots, on years 2, 6, 12, 17, and 22 (Hastie & Tibshirani, 1990).

After restructuring the dataset into person-period format, we divided the dataset into two mutually exclusive and exhaustive subsets, the first dataset included all the person-periods up to time 4 (1992) and the second dataset included all the person-periods with a time period greater than 4. Considering that all observations had either an event (all-cause mortality) or were at least observed for 3 years, the first dataset was used for learning and prediction model building to obtain the predicted survival outcome for those observations at time 4. The second dataset was restricted to the observations that were successfully followed at the Phase IV of the study [all with event or censoring time greater than 4] and was used to build a prediction model and obtain survival outcome prediction for those observations censored before the administrative end of the follow-up (end of 2011). The second dataset was weighted to be representative of the initial full sample for which the estimators will be used to obtain prediction of the survival outcomes. The appendix section titled observation weights for the second sub-sample used in the Super Learner, briefly describes the calculation of the time dependent weights (inverse probability of non-censoring weights) for this sub-sample.

We implemented super learning using the SuperLearner package in R (E.C. Polley, LeDell, & van der Laan, 2016). The ensemble library for the super learner for both datasets included the adaptive splines algorithm (i.e. earth package in R), penalized regression using elastic net (i.e. glmnet package in R), random forest (i.e. randomForest package in R), Bayesian generalized linear regression, linear discriminant analysis, and generalized linear models. Since individuals had multiple observations, one for each discrete time, for creating the V-fold cross validation subsets we considered the participants' unique identifier (ID) as a clustering variable to force the observations with the same ID in the same sub-sample. We used 10-fold cross-validation for

this learning. The parameters for a few of the algorithms were tuned to create competing algorithms and were all included in the ensemble library. The event indicator (all-cause mortality) at each discrete time was used as the outcome of interest. The predictor variables included in the dataset were the discrete time and its corresponding splines variables, age at the time of the earthquake and its square, gender, years of formal education, smoking during the earthquake, exercise during the earthquake, alcohol consumption frequency during the earthquake, weight at the time of the earthquake, height at the time of the earthquake, self-reported living standard, place of residence (region), type of residence (urban vs. rural), death of a nuclear family member due to earthquake, earthquake-related injury, loss of housing due to earthquake, job loss after the earthquake, receiving a permanent housing after the earthquake, and a set of indicator variables on various health conditions (i.e. cardiovascular diseases and hypertension, diabetes mellitus, stroke, asthma, chronic respiratory diseases, gastrointestinal tract diseases, urinary tract and kidney problems, arthritis, migraine, allergy, cancer, and mental health problems). The super learner was specified to minimize the rank loss function, maximizing the AUROC (a function of specificity and sensitivity in the validation sample).

The super learner prediction algorithm was used to estimate the probability of the outcome (death) at each time point (hazard) within the time period not observed because of censoring. We created 24 copies of the person-oriented dataset that included all eligible participants. The copies were indexed from 1 to 24. The observations in the first 23 copies were assigned to have the event at the time corresponding to their index, 1 through 23. The observations indexed 24 were assigned to be event free and were assumed to be administratively censored at the end of year 23.

Using the super learner algorithm the estimated probability of the outcome (an event) at each time point (hazard) during the time periods not observed because of censoring was obtained

and used to create weights for the above described 24 copies of the dataset. The weight of each participant in each copy of the dataset corresponded to the difference between its cumulative survival probability at the end of pervious time and its cumulative survival probability at its indexed time point. The observations in the last copy of the dataset (indexed 24) were assumed to be censored and were weighted by the complement of the cumulative survival probability at the end of year 23. The sum of the weight of all datelines (24 copy) for each individual was 1.The weighted person-oriented dataset was used in the analyses described in Chapter 3.

Random survival forests

Random survival forests (Ishwaran, Kogalur, Blackstone, & Lauer, 2008) is an ensemble tree method for analyzing time-to-event data, including right-censored data. It is based on the random forests method, an ensemble method for classification and regression problems that is boosted using randomization in the learning process (Breiman, 2001). The randomization is utilized in two steps in the process: the first is the random selection of a bootstrap sample, from the original sample, to serve as the learning set for growing trees; and the second is the random selection of a subset of candidate variables for splitting the selected observations at each node.

The random survival forests algorithm draws *B* bootstrap samples from the original study sample. On average, each bootstrap sample excludes one third (about 37%) of the observations, which later is used for the out-of-bag error rate calculations. The algorithm grows a survival tree based on each of the bootstrap samples. Each node of the tree splits the observations by selecting *p* candidate variables from the universe of variables available in the original dataset. The growing process starts with each bootstrap sample and each node is split into two daughter nodes. The purpose of the splitting is to maximize the survival difference

between the daughter nodes by examining and selecting the variables and the split values for each variable at each node. The algorithm offers a few splitting rules that can be used to maximize the differences between the observations. Further discussion of the splitting rules are provided here (Ishwaran & Kogalur, 2007). Each new node is further split into two new nodes until a node is saturated, a point is reached where the daughter nodes do not meet the requirement of having a minimum d_0 unique events (in our case deaths) [d_0 must be greater than 0]. Subsequently, the cumulative hazard function from each tree is the Nelson-Aalen estimator and the average of these cumulative hazard functions yields the ensemble cumulative hazard function. The observations not included in the bootstrap sample are used to estimate the C-index (concordance index) and calculate the out-of-bag error rate (prediction error) for the ensemble cumulative hazard function.

Similar to the random forests, the random survival forests method offers several advantages. Notably, the ease of use, as there are only three parameters that analysts need to set, namely the number of trees, the number of randomly selected predictors, and the splitting rule). Moreover, the method is data adaptive and does not impose restricting model assumptions (such as the proportional hazard assumption when dealing with time-to-event data). Further details about the random survival forests theory and its application are provided here (Ishwaran & Kogalur, 2007; Ishwaran et al., 2008; Mogensen, Ishwaran, & Gerds, 2012).

We used the randomForestSRC package in R to implement the random survival forest algorithm and obtain the cumulative hazard function in our data and apply that function to predict the survival probability of the censored observations up to the administrative end of the follow-up period. We split the existing observations into two exclusive dataset, one containing all the observations where those with greater than 4 years of follow-up time were artificially censored at the end of the 4th year, and another dataset which only included those with follow-up time

greater than 4 years. The first dataset was used to obtain the cumulative hazard function for the first 4 years, while the second dataset was employed to obtain the cumulative hazard function for the time greater than 4 years. Both datasets were transformed into the person-oriented format, where each participant had one line of observation with an indicator for event or censoring and a time variable. Since the second dataset included those observations that were successfully followed up to some point in time in the Phase IV, we weighted that sample to represent the original cohort. The weights were applied in the random survival forests when drawing the bootstrap samples. Details about estimation and creation of the weights are presented in the appendix. (Observation weights for the second sub-sample were used in the random survival forest algorithm.).

The random survival forest algorithms for both samples were administered and evaluated with a wide range of parameters for number of trees and number of candidate variables selected for splitting at each node. The potential predictors supplied to the random survival forest algorithm included age at the time of the earthquake and its square, gender, years of formal education, smoking during the earthquake, exercise during the earthquake, alcohol consumption frequency during the earthquake, weight at the time of the earthquake, height at the time of the earthquake, living standard, place of residence (region), type of residence (urban vs. rural), death of a nuclear family member due to earthquake, earthquake-related injury, loss of housing due to earthquake, job loss after the earthquake, and receiving a permanent housing after the earthquake. Variables on health profile of participants were also supplied to the algorithm, and those were a set of binary variables on cardiovascular diseases including hypertension, diabetes mellitus, stroke, asthma, chronic respiratory diseases, gastrointestinal tract diseases, urinary tract and kidney problems, arthritis, migraine, allergy, cancer, and mental health problems. All the analyses followed the log-rank splitting rule for splitting the observation at each node (Ishwaran & Kogalur, 2007; Leblanc & Crowley, 1993; Segal, 1988). The algorithms

were applied to obtain the survival probabilities for each participant in the original cohort. The probabilities were updated using the available information from the time period that participants were observed in the study. Twenty-four copies of the original person-oriented dataset (including all participants) were created, and the observations in each dataset were assigned 1 out of the 24 possible outcome scenarios. The observations were then weighted using the cumulative survival probabilities obtained from the random survival forest algorithms. The process of weighting was similar to the weighting described in greater detail under the SuperLearner section.

Multiple imputation (risk score imputation)

Multiple imputation is one of the most common and modern techniques for dealing with missing data in a wide range of analytical procedures. Nevertheless, due to the technical challenges, multiple imputation of time-to-event data is less commonly applied. Risk score imputation method is an approach for non-parametrically imputing time-to-event data using axillary variables, while relaxing the independent censoring assumption and recovering the information for those observations censored. The risk score imputation method uses the available auxiliary variables to calculate a risk score for each observation and define a set of nearest neighborhood observations for each observation (Hsu & Taylor, 2009).

This method fits two working proportional hazard models, one for the event time and another for the censoring time, both conditional on a set of user-selected auxiliary variables. The risk scores obtained from these two working models are further standardized to have a mean of zero and a standard deviation of 1. These two scale-free (standardized) risk scores along with corresponding user defined weights are used to obtain the distance between the observations. The risk set for subject *i* with a non-administrative censoring time, *t*, consists of *n* subjects

(value of *n* is specified by the analyst) that have censoring or event time that is greater than *t* and the *n* smallest distance from the subject *i*. If the number of available subjects with a censoring or event time greater than *t*, *m*, is smaller than *n*, then the size of the risk set for that individual will be *m* instead of *n*. The observations in each risk set are used to develop a Kaplan-Meier estimator. For each observation a random draw from $U \sim [0, 1]$ is made. The time corresponding to the Kaplan-Meier estimator at *U* is taken as the imputed time. The U is also used to impute an event indicator (where 1 stands for event and 0 for censoring) with $P(event=1)=U$. The process is repeated for each non-administratively censored observation to impute the time and event indicator. The risk scores for each imputation set are obtained from fitting two working proportional hazard models (for event time and for censoring time) on a bootstrap sample drawn from the original sample.

We implemented the risk score multiple imputation using the InformativeCensoring package in R. The risk set and the number of imputations were specified to be 10 and 20 respectively. The model for the event (death) time included age, age squared, gender, place of residence (region), type of residence (urban vs. rural), housing damage, death of a family member, serious injury, and receiving permanent housing after the earthquake. The model also included an interaction term between housing damage and receiving housing after the earthquake. The model for censoring time included age, gender, region, living standard before earthquake, housing damage, receiving housing after the earthquake and death of a family member. The final imputed datasets served as a source for sensitivity analyses presented in Chapter 3.

Statistical analysis

The statistical analysis approach for each research question is described in its corresponding chapter. Here we provide a summary and a discussion of mediation analyses in causal

inference with the intention to provide a background for the mediation analysis framework and justification for the specific types of effect decomposition and analytical methods employed in Chapters 3 and 4.

2.3 Mediation analysis in causal inference

Introduction to mediation analysis

The introduction of identification and estimation methods during the last few decades has resulted in rapid expansion of the causal inference literature. Such developments have contributed to the growth in number and improvement in quality of studies investigating causal relationships.

Evaluation of theories about and explanation of potential mechanisms for exposure effect on the outcome of interest is fundamental to nearly all scientific research. Mediation analysis is a tool that enables researchers to examine potential mechanisms for causal effects (Kraemer, Stice, Kazdin, Offord, & Kupfer, 2001) and it allows for explanation and better understanding of the nature of causal effect of the exposure on the outcome (Hafeman & Schwartz, 2009; Kenny, 2008). Hence, mediation analysis is an important and vital topic in many disciplines.

Mediation analysis can be used to test mediating pathways and mechanisms in both experimental and observational research (MacKinnon et al., 2013) and allows for empirical evaluation of theories and hypotheses; by doing so, it provides a ground for causal inference. For instance, identification of a hypothesized mediator can provide supporting evidence for existence of a causal relation between the observed exposure and outcome (Hafeman & Schwartz, 2009). Mediation analysis also enables researchers to distinguish true null effects from non-null pathway specific effects averaging to a total null effect. In extreme cases one or a

few of pathway-specific effects can be in the opposite direction of the total effect of the exposure on the outcome.

Often scientific literature do not provide a useful guidance for policy level changes and implementation of interventions (Judea Pearl, 2014b). Mediation analysis framework can offer more accessible and relevant information for policymaking and by doing so address the gap between science and policy. Additionally, in some circumstances it might not be feasible or ethical to directly intervene on the exposure, in such situations, knowledge about mediating mechanisms between the exposure and outcome can provide alternative options for targeting the causal effect of the exposure on the outcome. Furthermore, mediation analysis can be used to evaluate interventions and quantify the potential impact of various interventions through different mechanisms which in turn can be used to optimize existing interventions (Vanderweele, 2015).

The history of mediation dates back to early 1900s (Wright, 1923). Wright was one of the first to introduce the mediation concept into the statistics literature (Wright, 1934). Similarly, Ronald Fisher was among the front runners of mediation in the experimental design literature (Fisher, 1935). Social scientist, Hymen, in his survey design and analysis book suggested a three step strategy for mediation analysis, calling it ‘elaboration’ (Hyman, 1955; Kenny, 2008). In a nutshell, those three steps include 1) testing the association of exposure and outcome without including the candidate mediator in the model, 2) testing the association between the candidate mediator and the outcome of interest, and 3) Showing that adjusting for the mediator decreases the magnitude of the association between exposure and outcome. Since then, mediation analysis, like many other fields, has undergone significant developments. The next section highlights the traditional approaches employed for mediation analysis and the potential limitations inherited to those approaches.

Traditional approaches to mediation analysis

Most of mediation analysis techniques have roots in path analysis methods (Wright, 1934). For the discussion of mediation analysis approaches, we will consider the following data generating mechanism in a linear system, where *X* is the exposure of interest, *Y*, the outcome, and *M*, a mediator of the effect of the exposure on the outcome.

$$x = \alpha_0 + \varepsilon_x \quad (2.1)$$

$$m = \beta_0 + \beta_1 x + \varepsilon_m \quad (2.2)$$

$$y = \theta_0 + \theta_1 x + \theta_2 m + \varepsilon_y \quad (2.3)$$

In the social sciences, particularly in psychology, a majority of studies implementing mediation analysis have applied structural equation modeling (SEM) techniques. Product method (Baron & Kenny, 1986; Hyman, 1955; Sobel, 1982) and difference methods are two commonly used approaches that incorporate SEM. In the product method, the idea is to estimate the effect of *X* on *M* and also the effect of *M* on the outcome, *Y*. The effect of *X* on *M* is represented by β_1 from equation 2.2 and the effect of *M* on *Y* is θ_2 from equation 2.3. Simply, $\beta_1 * \theta_2$ would be the product method estimator for the effect of *X* on *Y* that is mediated through *M*.

The difference method obtains the indirect and the direct effects of *X* on the outcome by fitting two outcome regression models. One with the exposure and any potential confounders for the total effect of *X* on *Y* and another with the addition of the mediator and mediator-outcome confounders. The coefficient of *X* in the later model (2.3) would provide an estimate of the direct effect of *X* on *Y*, and the difference between coefficients of exposure in equations 2.2 and 2.3 will represent the indirect effect of *X* on the outcome that is mediated through *M*. The advantages of the above suggested methods are their simplicity to implement and understand.

However, these methods do not allow for interaction between the exposure and mediator. Moreover, in nonlinear settings, such as logistic regression, and when the outcome is not rare, the product and difference methods provide different estimates, where none has a causal interpretation. The issue arises because of non-collapsibility of the effects measures obtained from nonlinear models (Greenland, Robins, & Pearl, 1999). Nonetheless, it is worth noting that the product method can still be a valid tool to assess the existence of mediation but not estimation of an effect (Vanderweele, 2015).

Approaches that allow for mediation analysis in the presence of exposure-mediator interaction have been proposed (e.g.(Preacher, Rucker, & Hayes, 2007)). For instance, modifying the outcome model as in equation 2.4 allows for an interaction between the exposure and the mediator.

$$E(Y|m,x) = \lambda_0 + \lambda_1 x + \lambda_2 m + \lambda_3 xm \qquad (2.4)$$

The point estimate of the conditional indirect effect can be defined using the following $\beta_1\ (\lambda_2 + \lambda_3 m) * x$. Nevertheless, the direct and indirect effects defined in this method often do not sum up to the total effect, hindering estimation of the path specific contribution of the exposure on the outcome (Valeri & Vanderweele, 2013).

Researchers have successfully introduced different approaches that address the limitations of the traditional mediation analysis approaches and allow for interaction between the exposure and mediator while extending the mediation analysis to nonlinear equations/models (Halpern, 2000; Holland, 1988; Judea Pearl, 2014a; Robins & Greenland, 1992). The counterfactual framework (Rubin, 2005) is a flexible approach that addresses the limitations faced by the traditional mediation analysis and can offer accessible definition for various direct and indirect

effects and guide the estimation of their empirical analogous. Below we define several direct and indirect effects using the potential outcome (counterfactual) framework (Rubin, 2005) . Later, we also provide a summary of the required assumptions for causal mediation analysis.

Potential outcomes framework for mediation analysis

Notation

Let us assume the causal representation based on the equations 2.1, 2.2, and 2.4, where *X* is the exposure of interest, *M* is a mediator, and *Y* the outcome. Let Y_x represent the potential outcome that would have been observed had *X* been set to *x*. Similarly, M_x would be the potential value of mediator that would have been observed had *X* been set to *x*, while Y_{xm} would be the potential value when *X* and *M* are set to *x* and *m*. The potential outcomes can be used to represent exposure intervention specific outcomes, as well as nested counterfactuals indicating the 'natural' mediator effect on the potential outcome. For instance $Y_{xM_{x^*}}$ would represent the outcome value under *X* set to its reference value x^* and *M* being set to a value that would have been observed under setting *X* to *x*. For the pedagogical purposes and simplicity of this example, in this section we will assume all variables are binary with 1 representing the index and 0 the reference levels.

In practice, we cannot observe more than one potential outcome for a given subject. Hence, generally it is not practical to estimate a subject specific causal effect. Nonetheless, under certain assumptions we can define the average causal effects for a population or a sub-population. The potential outcomes framework provides a compact way of representing various causal contrasts. The average total effect (TE) of *X* on *Y* is defined as the contrast of potential outcomes under *X*=1 and *X*=0 respectively and can be represented as $E(Y_1 - Y_0)$. Note that the

total effect does not condition on the distribution of the mediator, and it allows it to obtain its natural value under corresponding exposure level.

Mediation analysis allows decomposition of the total effect into different components. There are various ways of effect decomposition. In this section we will cover the 4-way-effect decomposition followed by the interventional approach for path-specific effect decomposition.

4-way-effect decomposition

Recently, VanderWeele (2014) introduced a 4-way-effect decomposition technique that splits the total effect into the controlled direct effect (CDE), the reference interaction effect (RIE), the mediated interaction effect (MIE), and the pure indirect effect (PIE). Figure 2.2 presents various types of 2-way, 3-way and 4-way effect decomposition and their components.

	TE: $(Y_x - Y_{x^*})$			
2 way	PDE: $(Y_{xM_{x^*}} - Y_{x^*M_{x^*}})$		TIE: $(Y_{xM_x} - Y_{x\,M_{x^*}})$	
2 way	TDE: $(Y_{xM_x} - Y_{x^*M_x})$			PIE: $(Y_{x^*M_x} - Y_{x^*M_{x^*}})$
3 way	CDE: $(Y_{xm^*} - Y_{x^*m^*})$	PAI: $(Y_{xm} - Y_{xm^*} - Y_{x^*m} + Y_{x^*m^*})1(\mathrm{M}_x)$		PIE: $(Y_{x^*M_x} - Y_{x^*M_{x^*}})$
3 way	PDE: $(Y_{xM_{x^*}} - Y_{x^*M_{x^*}})$		MIE: $(Y_{xM_x} - Y_{x\,M_{x^*}} - Y_{x^*M_x} + Y_{x^*M_{x^*}})$	PIE: $(Y_{x^*M_x} - Y_{x^*M_{x^*}})$
4 way	CDE: $(Y_{xm^*} - Y_{x^*m^*})$	RIE: $(Y_{xm} - Y_{xm^*} - Y_{x^*m} + Y_{x^*m^*})1(\mathrm{M}_{x^*})$	MIE: $(Y_{xM_x} - Y_{x\,M_{x^*}} - Y_{x^*M_x} + Y_{x^*M_{x^*}})$	PIE: $(Y_{x^*M_x} - Y_{x^*M_{x^*}})$

Figure 2.2 Summary of various effect decomposition in causal mediation analysis

TE: Total effect, PDE: pure direct effect, TIE: total indirect effect, TDE: total direct effect, CDE: controlled direct effect, NDE: natural direct effect, PAI: portion attributable to interaction, PIE: pure indirect effect, MIE: mediated interaction effect, RIE: reference interaction effect.

The CDE is a contrast between the potential outcomes setting *X*=1 and *X*=0 while the mediator *M* is fixed to a specific level *m* and can be represented as $E(Y_{1m} - Y_{0m})$. In estimating CDE, since there is no restriction on the level of M that the population is fixed to. In the presence of exposure-mediator (X-M) interaction, the estimated CDE can vary by the choice of *m*. The natural direct effect (NDE) is another measure of direct effect which fixes the mediator to a specific natural *m* that each individual level would have experienced under either X=1 or X=0; thus, unlike CDE, it allows for variations in *M* at the population level. The NDE compares potential outcomes under *X*=1 and *X*=0 while the mediator M is fixed to the level (M_0) that would have been observed under reference level of *X*. In the potential outcome framework NDE is defined as $E(Y_{1M_0} - Y_{0M_0})$. This quantity is sometimes referred as pure direct effect (PDE). The natural indirect effect (NIE), also known as the average total indirect effect (TIE), is the contrast between the outcomes under M_x and M_{x^*} while setting exposure *X* to its index value *X*=1 (more generally, *X*=*x* is the index exposure value while *X*=*x** is the reference exposure value).

Components of various effect decompositions can be obtained through combination of different component effects. For instance, a method of 2-way decomposition combines the CDE and RIE into the PDE. The remaining components are summed into TIE.

Assumptions for identification and estimation of effects in mediation analysis

For the estimation and identification of effects we make the following assumptions (Wang & Arah, 2015)

1. Conditional exchangeability. This translates into the causal assumption of no uncontrolled confounding conditional on the set of covariates (Z). For the mediation analysis, these assumptions include (J Pearl, 2001; Robins & Greenland, 1992; Vanderweele, Vansteelandt, & Robins, 2014)

a) No uncontrolled confounding of exposure-outcome relation conditional on the set of measured covariates, *Z*.
b) No uncontrolled confounding of mediator-outcome relation conditional on the exposure, *X*, and the set of measured covariates, *Z*.
c) No uncontrolled confounding of exposure-mediator relation conditional on the set of measured covariates, *Z*.
d) No mediator-outcome confounder is affected by the exposure.

This is a comprehensive list of exchangeability assumptions for mediation analysis in general. Identification of some of the effect components might only require a subset of the assumptions.

2. Stable unit treatment value assumption (SUTVA) (Robins, 1986)
3. Positivity (Hernán & Robins, 2006)
4. Consistency (Judea Pearl, 2010; Robins, Hernán, & Brumback, 2000)

Despite the growing literature on causal inference and popularity of mediation analysis, most of the methods on mediation analysis have focused on settings with a single mediator and confined to certain types of outcomes. Majority of the methods also fail to formally represent the assumptions for the causal mediation analysis. Moreover, only a few of the methods are applicable to settings where the no exposure induced mediator outcome confounding assumption can be relaxed. Although the 4-way decomposition is quite flexible framework and the estimated effect component can be used to construct several other types of effect decomposition including natural (in)direct effects, it requires the no-exposure induced mediator-outcome confounding strong assumption. That assumption is not always warranted, and in settings with multiple mediators it is quite often violated. The mediation analysis using

interventional approach for path-specific effect estimation is a mediation analysis framework that overcomes that shortcoming and relaxes that assumption.

Interventional approach for path-specific effect estimation

The interventional natural direct and indirect effects framework proposed and introduced by VanderWeele et al. (2014) overcomes the need for conditional independence of the cross-word counterfactuals and makes the estimation task possible under weaker assumptions at the expense of obtaining the natural decomposition originally introduced by Robins, Greenland and Pearl. Nevertheless, the interventional natural direct and indirect effects do not always add up to the total effect. Vansteelandt and Daniel (2017) adapted the VanderWeele et al. (2014) framework to overcome the drawback of not having the effect components always adding up to the total effect. The modified strategy, interventional approach for path-specific effects estimation, can be extended to the multiple-mediator settings where the total effect is decomposed into path specific effects, even when causal structure between the mediators is unknown.

To provide a discussion on identification and estimation of interventional approach for path-specific effects and go through the steps, we assume the causal structure depicted in figure 2.3, where X is the exposure of interest, M_1 and M_2 the mediators, Y the outcome, C a vector of measured covariates, and U a vector of unmeasured covariates.

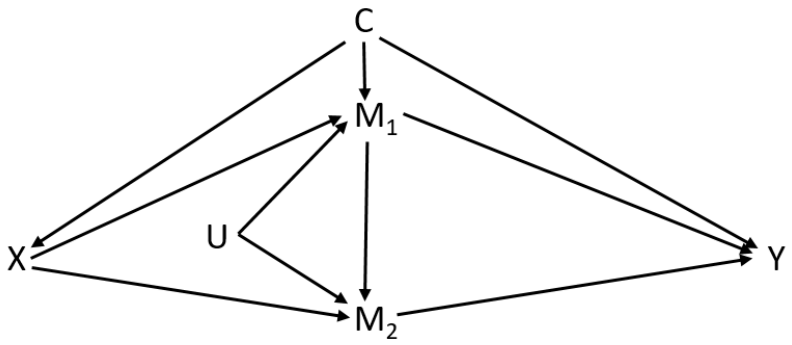

Figure 2.3 Causal diagram; X is the exposure of interest, M1 and M2 the mediators, Y the outcome, C a vector of measured covariates, and U a vector of unmeasured.

In a setting with two mediators, the interventional path specific approach decomposes the total effect into 4 components.

1) The interventional direct effect is defined as $E(Y_{x=1\ M_{1\ x=0}\ M_{2\ x=0}} - Y_{x=0\ M_{1\ x=0}\ M_{2\ x=0}})$
2) Interventional effect through M_1 is defined as $E(Y_{x=1\ M_{1\ x=1}\ M_{2\ x=0}} - Y_{x=1\ M_{1\ x=0}\ M_{2\ x=0}})$
3) Interventional effect through M_2 is defined as $E(Y_{x=1\ M_{1\ x=0}\ M_{2\ x=1}} - Y_{x=1\ M_{1\ x=0}\ M_{2\ x=0}})$
4) Interventional effect mediated due to dependence of M_1 and M_2 is defined as

$$E\left(Y_{x=1\ M_{1\ x=1}\ M_{2\ x=1}} - Y_{x=1\ M_{1\ x=0}\ M_{2\ x=1}}\right) - E\left(Y_{x=1\ M_{1\ x=1}\ M_{2\ x=0}} - Y_{x=1\ M_{1\ x=0}\ M_{2\ x=0}}\right)$$

This brief review about selected topics in mediation analysis was intended to provide a background and rationale for some of the decisions and analytical approaches that will be considered while addressing the research questions in the upcoming chapters (Chapters 3 and 4).

2.4 Appendix

Observation weights for the second sub-sample used in the Super Learner

The weight for the subset of observations of person-period data with event or censoring time greater than 4 were calculated using the data from the full sample, including those who had a censoring time or event time of less than 4. A censoring indicator variable, C, was created to represent if the person was censored at the corresponding time in the dataset. All the observations with non-administrative censoring were assigned C=1 at the time censored, while all others were assigned C=0. The numerator for the weight was equal to the probability of not being censored at the corresponding time in the dataset and the denominator was estimated to be equal to the probability of not being censored at the corresponding time conditional on baseline characteristics, *Z*, (age, age squared, gender, place of residence (region), housing damage, death of a nuclear family member due to earthquake, job loss after the earthquake, and receiving a permanent housing after the earthquake).

The final weights for subjects who were at risk of the event and uncensored at time *t* was defined as CW(t) where

$$CW(t) = \prod_{k=1}^{t} \frac{\Pr[\mathrm{C(k)} = 0]}{\Pr[\mathrm{C(k)} = 0 | \mathrm{Z} = \mathrm{z}]}$$

There are examples of studies describing, applying, and providing code for censoring weights (Hernán, Brumback, & Robins, 2000, 2001)

Observation weights for the second sub-sample used in the random survival forest algorithm

The weights for the subset of observations in the person-period data with an event or censoring time greater than 4 were calculated using the data from the full sample, including those who had a censoring time or event time of less than 4. A selection indicator variable, S, was created to represent if the person had an event (at any time) or had a censoring time greater than 4. The numerator of the weight was estimated to be equal to the probability of selection and the denominator was the probability of selection conditional on the baseline characteristics, (age, age squared, gender, place of residence (region), housing damage, job loss after the earthquake, death of a nuclear family member due to earthquake and receiving a permanent housing after the earthquake), C.

The final weights for subjects who were alive and uncensored at the end of time 4 were defined as SW where

$$SW = \frac{\Pr[\mathrm{S} = 0]}{\Pr[\mathrm{S} = 0|\mathrm{C} = \mathrm{c}]}$$

Chapter 3. Impact of earthquake exposure on and the mediating role of post-earthquake experiences in all-cause mortality

Abstract

Despite the substantial negative health and social impact of earthquakes, studies exploring the long-term survival in the aftermath of earthquakes are scarce. The current study used data from a prospective cohort of surviving adults with differential exposure levels to the 1988 earthquake in Armenia to investigate the impact of earthquake-related exposure (i.e. housing damage, death of a family member due to the earthquake, and severe earthquake-related injury) on all-cause mortality. A total of 23,639 individuals were followed after the earthquake for at least 3 years, with a sub-cohort of more than 1700 individuals followed up to 23 years. Data adaptive and multiple imputation methods were employed to impute the survival outcome of censored individuals. We transformed the dataset with the imputed outcomes into discrete person-time and created natural cubic splines of time since the earthquake, with five knots at years 2, 6, 12, 17, and 22. The discrete time survival analysis adjusting for age, gender, education, standard of living, place of residence, earthquake related death in the family, and injury was used to obtain the hazard of death for individuals with different levels of housing damage. A Monte Carlo simulation was applied to integrate the hazards over time and estimate the cumulative survival probabilities. Finally, the effect of housing damage on all-cause mortality was decomposed with respect to receiving permanent housing in the aftermath of the earthquake. Those with no damage to their housing due to the earthquake had an average 23-year cumulative survival probability of 0.744 (95%CI: 0.735, 0.752), while those whose housing was completely destroyed had a 23-year cumulative survival probability of 0.721 (95%CI: 0.709, 0.733). Death of a family member due to the earthquake did not appear to be associated with all-cause mortality during the 23-year follow-up period. Those who sustained severe earthquake-related

injury and those without serious injury showed no differences during the first 10 years of follow-up. However, during the second decade of follow-up, the survival curves slowly diverged and the difference in cumulative survival probability of those with severe injury as compared to those without such injury reached −0.033 (95%CI: −0.056, −0.009). The effect of housing damage on all-cause mortality was partially mediated (mitigated) by receiving housing aid after the earthquake. Our results indicate that total destruction of housing had a noticeable and long-lasting effect on survival. The findings could help public health practitioners target vulnerable populations and guide future research.

3.1 Introduction

Natural disasters are inevitable and common events that can lead to a variety of short- and long-term physical, mental, social, and financial difficulties (Armenian et al., 1998; Dell'Osso et al., 2011; A K Goenjian et al., 1995; Leor, Poole, & Kloner, 1996; Watanabe et al., 2008; Zwiebach, Rhodes, & Roemer, 2010). Earthquakes account for a notable proportion of all natural disasters and can result in severe environmental, economic, and health consequences.

Studies have shown that the severity of earthquake exposure is associated with negative health outcomes, including diabetes, cardiovascular diseases, depression, and PTSD (Armenian et al., 1998; Leor et al., 1996). Through their adverse health effects, earthquakes directly and indirectly, affect victims' quality of life and are considered a major public health concern (Khachadourian, Armenian, Demirchyan, & Goenjian, 2015). Loss of a family member to the earthquake, earthquake-related financial loss, loss of employment following an earthquake, and lack of post-earthquake support are among the factors contributing to earthquake-related health disparities (Armenian et al., 1998; Armenian, Noji, & Oganesian, 1992; Dell'OSso et al., 2011; Khachadourian et al., 2015).

Despite the significant negative health and social impact of earthquakes, to date, only a limited number of studies have explored the long-term survival among populations exposed to earthquakes (Ripoll Gallardo et al., 2018). Comparison of the number of sudden cardiac death among 1994 Northridge earthquake survivors with those occurring during the three-year period preceding the earthquake showed an immediate and acute increase in number of sudden cardiac death one day after the earthquakes and suggested that the psychological stress might have precipitated adverse health events among those predisposed to such events (Leor et al., 1996). Similarly, a study of myocardial infarction mortality after an earthquake in Japan found that populations living in the affected region had increased myocardial infarction related mortality during the 8-week period following the earthquake, and also showed that those areas with greater destruction of houses tended to have a higher myocardial infarction related mortality rates (Ogawa, Tsuji, Shiono, & Hisamichi, 2000).

Although the existing body of literature hints of increased mortality in the aftermath of disasters (Ripoll Gallardo et al., 2018), the methodological limitations observed in most such studies make their findings less conclusive. For instance, most studies offer a cross-sectional or time series design, or suffered from limited power (Dobson, Alexander, Malcolm, Steele, & Miles, 1991; Kario & Ohashi, 1997; Takegami et al., 2015; Yashiro et al., 2000). Use of aggregate exposure measures and lack of data on the individual level are among other important limitations faced by a large proportion of studies investigating the effects of earthquake on mortality. Additionally, some studies targeted a specialized study population or focused on specific types of mortality, and hence offer low generalizability to the often broad populations exposed to earthquakes. Moreover, most of previous studies report a follow-up time of less 5 years with the majority focusing on first few months to a year after the earthquake (Dobson et al., 1991; Kario & Ohashi, 1997; Nakagawa et al., 2009; Takegami et al., 2015; Yashiro et al., 2000).

The current study sought to investigate the impact of earthquake-related exposures and post-earthquake factors on all-cause mortality among a population exposed to the 1988 earthquake in Armenia.

The devastating earthquake that hit the northern part of Armenia in 1988, had a high death toll, and an immediate negative impact on life expectancy. In a prospective cohort study of the population exposed to the earthquake, Armenian et al. (1998) found increased rates of death from all-causes and especially from heart disease during the first year after the earthquake, which were followed by relatively lower and more stable mortality rates during the second year. Similar to the most of other populations exposed to earthquakes, the longer-term mortality outcomes of earthquake survivors in Armenia are unknown. The current study sought to assess the long-term impact of earthquake-related exposures and the mediating role of post-earthquake events on all-cause mortality among a cohort exposed to the 1988 Spitak earthquake in Armenia using data from the PEPSI. The specific research questions were as follow:

- What are the independent effects of earthquake-related exposures, namely death of a nuclear family member, serious injury, and loss of housing on all-cause mortality?
- What is the potential mediating or mitigating role of providing permanent housing in the potential effect of loss of housing on all-cause mortality?

3.2 Methods

Study sample and variables

We used data from PEPSI study, a prospective follow-up study of more than 30,000 individuals with differential exposure levels to the 1988 Spitak earthquake. The analyses were limited to

participants who were 18 or older at the time of the earthquake, and were either successfully followed for at least 3 years or had an event (death from any cause). Further description of the study design and methods is provided in Chapter 2 and elsewhere (Khachadourian et al., 2016).

Outcome

The survival outcome of individuals in the study was collected throughout the follow-up phases. Year of death and underlying cause of death were recorded for the deceased participants. The survival outcome of those participants censored before the administrative end of the follow-up was predicted and imputed using the Super Learner. Chapter 2 provides further details about obtaining the predicted probabilities of the outcome using the Super Learner.

Exposures

The housing damage variable was a categorical variable defined as a) no damage, b) minor damages/non-structural damages, and c) extensive damage/structural damages/total destruction. The number of nuclear family deaths caused by the earthquake was also collected. Since there were only a few individuals with multiple deaths in the family, the variable was dichotomized to no death vs. any death. Details on earthquake-related injury collected at Phase I of the study included the body part affected and type of injury. We created a binary variable (yes/no) to summarize serious earthquake-related injury of the study participants. Scratches and minor cuts were not considered serious injuries.

Mediator (Post-earthquake variables)

Receiving permanent housing during the 3 years following the earthquake (measured during 1989-1991) was the mediator of interest.

Covariates

Data on age, gender, education, employment, pre-earthquake residence location and type, and living standard were collected at Phase I and were verified during the subsequent follow-up phases. Age was a continuous variable recorded in years. Education was recorded as a continuous variable reflecting years of formal education. Data on pre earthquake lifestyle measures including consumption of alcohol, smoking habits, and physical activity were also available and were considered in the analyses.

The original dataset included a total of 26,478 eligible participants. However, as described in Chapter 2, to assure full availability of data on mediator of interest during the first three years of the follow-up, all participants with a censoring time before the end of 1991 were excluded from the analytical sample. The final analytical sample totaled to 23,639.

Statistical analysis

Frequencies, proportions, means, and standard deviations were calculated to summarize the baseline characteristics of the study population by earthquake-related exposure variables. The analytical sample was weighted to represent the entire original sample, including those individuals who were adults at the time of the earthquake (18 and above) but were censored during the first three years of the follow-up.

To address the specific objectives of the study, we obtained the weighted person-oriented dataset that was created by the predicted survival outcomes estimated by the Super Learner algorithm (described in Chapter 2). We generated 5000 bootstrap samples where the probability of selection in a bootstrap sample for each observation was directly proportional to the product of the weight of the observation in the dataset with predicted outcomes using the Super Learner

and the weight of the observation from the dataset weighted to represent the entire original sample, including those individuals who were adults at the time of the earthquake (18 and above) but were censored during the first three years of the follow-up. We restructured all the person-oriented bootstrap samples into the person-period format, where every individual had one observation in the dataset for each discrete time.

Assuming consistency, uncontrolled confounding (conditional exchangeability), and positivity, we fit a pooled logistic regression to assess the independent impact of earthquake-related housing damage, earthquake-related injury, and earthquake related death in the nuclear family on all-cause mortality. Discrete time survival analysis can easily accommodate time varying effect estimates and relaxes the proportional hazard assumption inherited in regular Cox proportional hazard regression models. To estimate the time dependent intercepts, we used natural cubic splines of time since the earthquake [in years], with five knots (at years 2, 6, 12, 17, and 22) to create a smoothing function. The final model was adjusted for potential confounders of the earthquake-related exposures' effects on all-cause mortality, including age, age squared, gender, education in years, standard of living at the time of the earthquake, and place of residence (region).

For mediation analysis, we used the person-period bootstrap samples and followed a four-way decomposition framework to estimate the CDE, MIE, RIE, and PIE of housing damage with respect to post-earthquake provision of permanent housing on all-cause mortality. The definitions of the effects in the potential outcome framework are provided in Figure 2.2. We implemented g-computation methodology to create the potential outcomes and obtain the relevant contrasts and effect estimates. The outcome model included all the covariates, earthquake-related exposure variables, and the mediator of interest (receiving housing in the 3-year period after the earthquake). Similar to the model for the estimation of the total effects, the

mediation analysis model included interaction terms for exposure and mediator variables with the time spline variables, allowing the effect of exposure of interest and the mediator to vary by time.

The analyses were conducted in SAS 9.4. Parameter estimates from the statistical models described above combined with Monte Carlo simulations were the basis for integrating the estimated hazards, estimating the cumulative survival probabilities and obtaining the corresponding effect measures. The final effect measures on the absolute scale (i.e. survival probability difference) were reported for year 1, 2, 3, 5, 10, and 23 of the follow-up. The final point estimates of the effect measures of interest and their 95% confidence intervals (95% CIs) were based on the corresponding percentile of the ordered estimates obtained from the 5000 bootstrap samples.

Sensitivity analysis

To assess the robustness of our findings regarding the effect of earthquake-related exposure on all-cause mortality, we repeated the main analyses while further adjusting for smoking and alcohol consumption habits of participants at the baseline. We also replicated the main analysis using datasets with outcome data, partially predicted and imputed by random survival forest as well as a score based multiple imputation algorithms.

The sensitivity of results of the mediation analyses from decomposition of the total effect of housing damage on all-cause mortality with respect to post-earthquake housing aid were evaluated through adjustment for smoking and alcohol consumption. We also repeated the mediation analyses using the datasets with outcome data partially imputed by random survival forest as well as a score based multiple imputation algorithms.

3.3 Results

Characteristics of the study sample

The total analytical sample included 23,639 participants. After applying the normalized weights, participants had an average age of 39.9 years. Slightly more than half of the sample (54.1%) were female. Participants had an average of 11.3 years of formal education at the time of the earthquake. About two-thirds of participants (65.3%) reported an average standard of living when compared to the general population, 30.5% reported below average standard of living, while a small fraction (4.2%) identified their standard of living in 1988 as above average. Slightly less than a third of the participants (30.1%) were current smokers at the time of the earthquake, while 13.3% reported exercising regularly before the earthquake. The majority of participants were living in urban areas at the time of the earthquake (83.5%) while others (16.5%) were residing in rural areas. Table 3.1 provides further details on characteristics of the weighted sample at the time of the earthquake.

Table 3.1 Baseline characteristics of the weighted sample at the time of the earthquake, in 1988

Characteristics	Total sample (N=23,639)
Age, years mean (SD)	39.9 (15.6)
Gender, n (%)	
Male	10841 (45.9)
Female	12798 (54.1)
Education, years mean (SD)	11.3 (3.2)
Living standard in 1988, n (%)	
Above average	985 (4.2)
Average	15445 (65.3)
Below average	7209 (30.5)
Smoking status in 1988, n (%)	
Smoker	7123 (30.1)
Non-smoker	16516 (69.9)
Drinking in 1988, n (%)	
Yes	7728 (32.7)
No	15911 (67.3)
Regular exercise, n (%)	
Yes	3157 (13.3)
No	20482 (86.7)
Area of residence, n (%)	
Spitak	1374 (5.8)
Gyumri	8454 (35.8)
Vanadzor	5923 (25.1)
Other areas	7888 (33.4)
Type of residence, n (%)	
Rural	3892 (16.5)
Urban	19747 (83.5)

Assessment of the earthquake-related exposures and post-earthquake experiences by region showed that those participants living in the Spitak had the highest levels of damages. Four out of five (79.5%) in Spitak had a total destruction of housing due to the earthquake. The earthquake completely destroyed the housing of 24.8%, 7.1%, and 12.1% of participants living in Gyumri, Vanadzor, and other regions within the disaster area respectively. More than a third of the participants (39.4%) in Spitak lost a nuclear family member to the earthquake, while the prevalence of earthquake-related death of a nuclear family member in Gyumri was 16.1%, in

Vanadzor 0.4%, and in other regions (moderately affected by the earthquake) 0.9%. Similar to other earthquake-related exposures, earthquake-related injury was the highest in Spitak (9.5%), followed by Gyumri (5.3%), Vanadzor (2.5%), and other regions in the disaster area (2.1%). During the period from earthquake until the end of 1991, 22.5% of the participants lost their job and were unemployed for more than 6 months; the highest rates of job loss were reported in Spitak, followed by Gyumri. During the three calendar years after the earthquake, 11.6% of participants received permanent housing. Table 3.2 provides further details on the distribution of earthquake-related measures across the affected regions.

Table 3.2 Participants baseline characteristics by region of residence, at the time of the earthquake, December 1988

Earthquake-related exposure and post-earthquake events	Spitak	Gyumri	Vanadzor	Other[a]	Total
Housing damage, n (%)					
No damage	49 (3.5)	3057 (36.2)	2409 (40.7)	3085 (39.1)	8601 (36.4)
Moderate damage	233 (17.0)	3300 (39.0)	3094 (52.2)	3850 (48.8)	10477 (44.3)
Total destruction	1092 (79.5)	2096 (24.8)	420 (7.1)	952 (12.1)	4561 (19.3)
Earthquake-related death in the nuclear family, n (%)					
Yes	586 (39.4)	1393 (16.1)	24 (0.4)	73 (0.9)	2075 (8.8)
No	902 (60.6)	7259 (83.9)	5762 (99.6)	7640 (99.1)	21562 (91.2)
Earthquake caused serious injury, n (%)					
Yes	131 (9.5)	429 (5.1)	148 (2.5)	167 (2.1)	875 (3.7)
No	1243 (90.5)	8025 (94.9)	5775 (97.5)	7721 (97.9)	22764 (96.3)
Job loss for more than 6 months during 1989-1991, n (%)					
Yes	491 (35.7)	2461 (29.1)	1206 (20.4)	1105 (14.0)	5263 (22.3)
No	883 (64.3)	5994 (70.9)	4717 (79.6)	6783 (86.0)	18376 (77.7)
Receiving a house as an aid during 1989-1991, n (%)					
Yes	204 (22.1)	1561 (18.5)	502 (8.5)	1188 (15.1)	3555 (15.0)
No	1070 (77.9)	6892 (81.5)	5421 (91.5)	6700 (84.9)	20084 (85.0)

[a] Other includes small cities and villages in the disaster affected areas

Effect of earthquake-related exposures on all-cause mortality

Housing damage

In the main analyses based on the data partially imputed using the Super Learner, the group without any earthquake damages to their housing had an average estimated 23-year cumulative survival probability of 0.755 (95%CI: 0.746, 0.765). The 23-year cumulative survival probability among those with moderate housing damage was 0.744 (95%CI: 0.735, 0.752), while those with total housing destruction of housing had a 23-year cumulative survival probability of 0.721 (95%CI: 0.709, 0.733). Figure 3.1 presents the cumulative survival probability of the study population over the follow-up period by the level of housing damage.

Those with a moderate damage of housing had a lower cumulative survival probability when compared to the population with no housing damage. The difference in cumulative survival probabilities across these two groups grew during the early years after the earthquake and then tended to be constant. The difference in cumulative survival probabilities at the end of the 23-year follow-up was −0.012 (95%CI: −0.023, −0.001). The difference in the cumulative survival probabilities of the population with total housing destruction compared to the population continuously increased over the entire study period, reaching a final difference in cumulative survival probabilities of −0.035 (95%CI: −0.049, −0.021) at the end of the follow-up period. Table 3.3 presents the marginal absolute effect estimates of the housing damage on cumulative survival probability at selected time points.

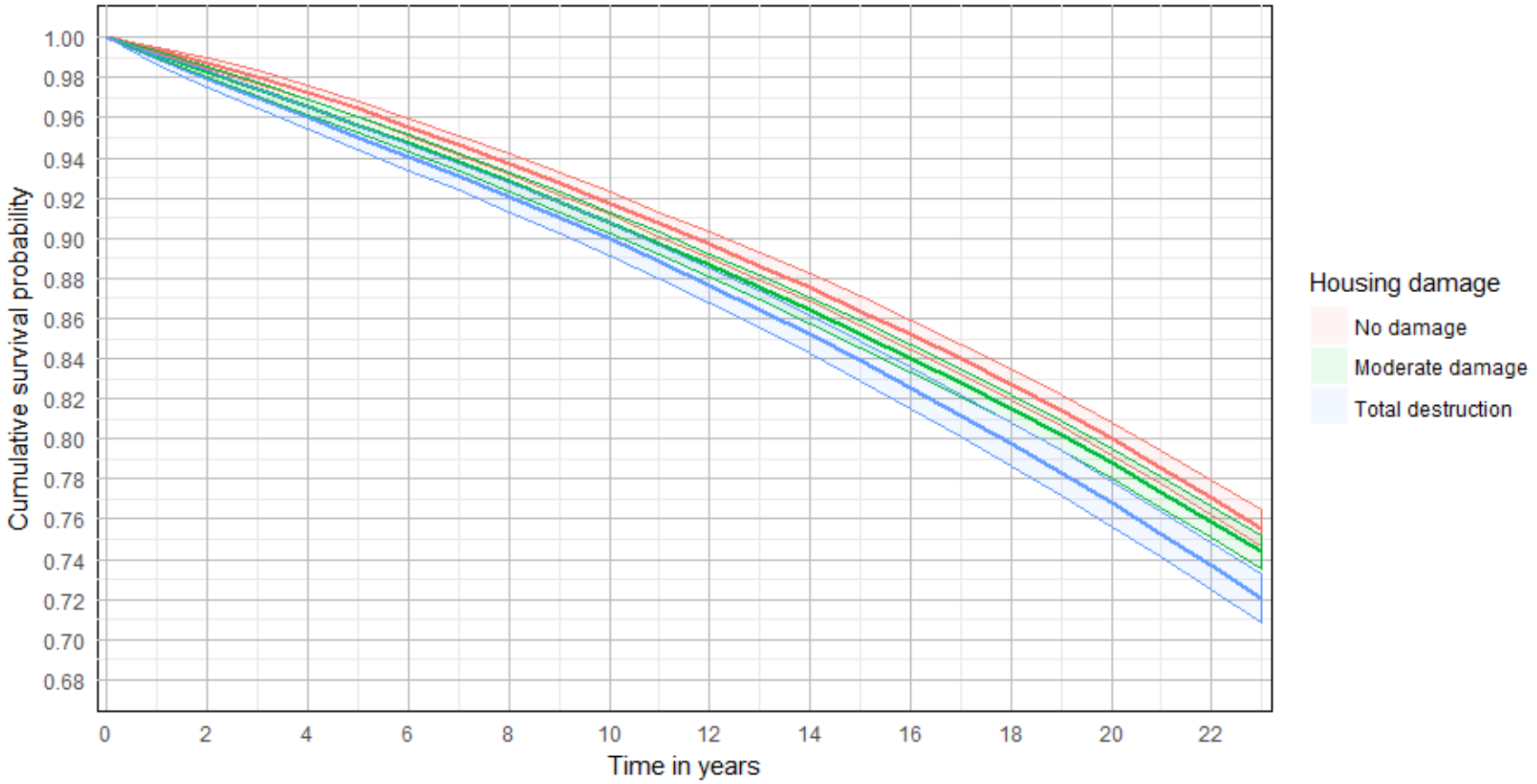

Figure 3.1 Cumulative survival probability among a cohort of the 1988 Spitak earthquake survivors during 1988-2012, by housing damage, using outcome data partially imputed by the Super Learner algorithm

[a] All the measures are standardized for age, age squared, gender, education in years, standard of living at the time of earthquake, place of residence (region), earthquake-related death of a family member, and serious injury among the total baseline population

[b] The point estimates and 95% confidence limits are based on the 50th, 2.5th, and 97.5th percentile obtained from 5000 bootstrap sample estimates

Table 3.3 Marginal absolute effects of earthquake-related housing damage on cumulative survival probability using outcome data partially imputed by the super learner algorithm among a cohort of the 1988 Spitak earthquake survivors

	Difference in cumulative survival probability [a] (95%CI) [b]		
	Housing damage		
Time (year)	**No damage**	**Moderate damage**	**Total destruction**
1	Reference	−0.003 (−0.005, 0.000)	−0.005 (−0.008, −0.002)
2	Reference	−0.005 (−0.008, −0.001)	−0.008 (−0.013, −0.003)
3	Reference	−0.006 (−0.010, −0.002)	−0.011 (−0.017, −0.005)
5	Reference	−0.008 (−0.013, −0.003)	−0.014 (−0.021, −0.007)
10	Reference	−0.010 (−0.017, −0.002)	−0.018 (−0.027, −0.009)
15	Reference	−0.012 (−0.020, −0.002)	−0.025 (−0.036, −0.013)
23	Reference	−0.012 (−0.023, −0.001)	−0.035 (−0.049, −0.021)

[a] All the measures are standardized for age, age squared, gender, education in years, standard of living at the time of earthquake, place of residence (region), earthquake-related death of a family member, and serious injury among the total baseline population

[b] The point estimates and 95% confidence limits are based on the 50th, 2.5th, and 97.5th percentile obtained from 5000 bootstrap sample estimates

Results of sensitivity analyses further adjusted for the smoking and drinking habits of the study participants at the baseline were consistent with the findings presented above. Table A-3.1 in the appendix presents the effect estimates from the sensitivity analyses.

Sensitivity analyses using data partially imputed through the random survival forests while otherwise following the main analyses approach yielded similar cumulative survival trends. Consistent with the findings from the main analyses, the populations with moderate housing damage and total destruction of housing had lower cumulative survival probabilities compared to the group who did not experience any housing damage. The overall survival curves and the effect estimates obtained using outcome data partially imputed through the random survival forest are presented in Figure A-3.1 and Table A-3.2 respectively.

Although the results of sensitivity analyses using outcome data partially imputed by the score-based multiple imputation algorithm to repeat the main analysis approach reiterated the negative impact of housing damage on cumulative survival probabilities, the effect estimates obtained from analysis of the data imputed using the score-based imputation algorithm tended to be moderately larger than those observed in the main analyses as well as the sensitivity analyses using the outcome data partially imputed through the random survival forests. Nevertheless, the confidence intervals across these analyses had notable overlaps and did not highlight any significant quantitative differences in the results. Figure A-3.2 and Table A-3.3 present overall survival curves and the effect estimates obtained using data partially imputed through the score-based multiple imputation algorithm respectively.

Death of a family member

The cumulative survival probabilities of those who lost a family member to the earthquake and those who did not experience such a loss did not differ over the follow-up time. Figure 3.2 provides the survival curves of the study population by earthquake-related death of a family member. Table 3.4 presents the marginal absolute effect estimates of loss of a family member to the earthquake on cumulative survival probability at selected time points over the study follow-up period. As depicted in Figure 3.2, at no point in time was the difference in the cumulative survival probabilities incompatible with the null effect.

In sensitivity analyses adjusting for smoking and alcohol consumption habit of study participants at the baseline, we did not observe any significant changes in the effect estimates of death of a family member on cumulative survival probabilities over time (Table A-3.4).

Further sensitivity analyses using outcome data partially imputed by random survival forest as well as using outcome data partially imputed with the score-based multiple imputation algorithm did not indicate any substantial changes when compared to the effect estimates from the main analysis. The cumulative survival probability plots, along with effect estimates from these analyses using different data sources, are presented in Figures A-3.3 and A-3.4, and Tables A-3.5 and A-3.6, respectively.

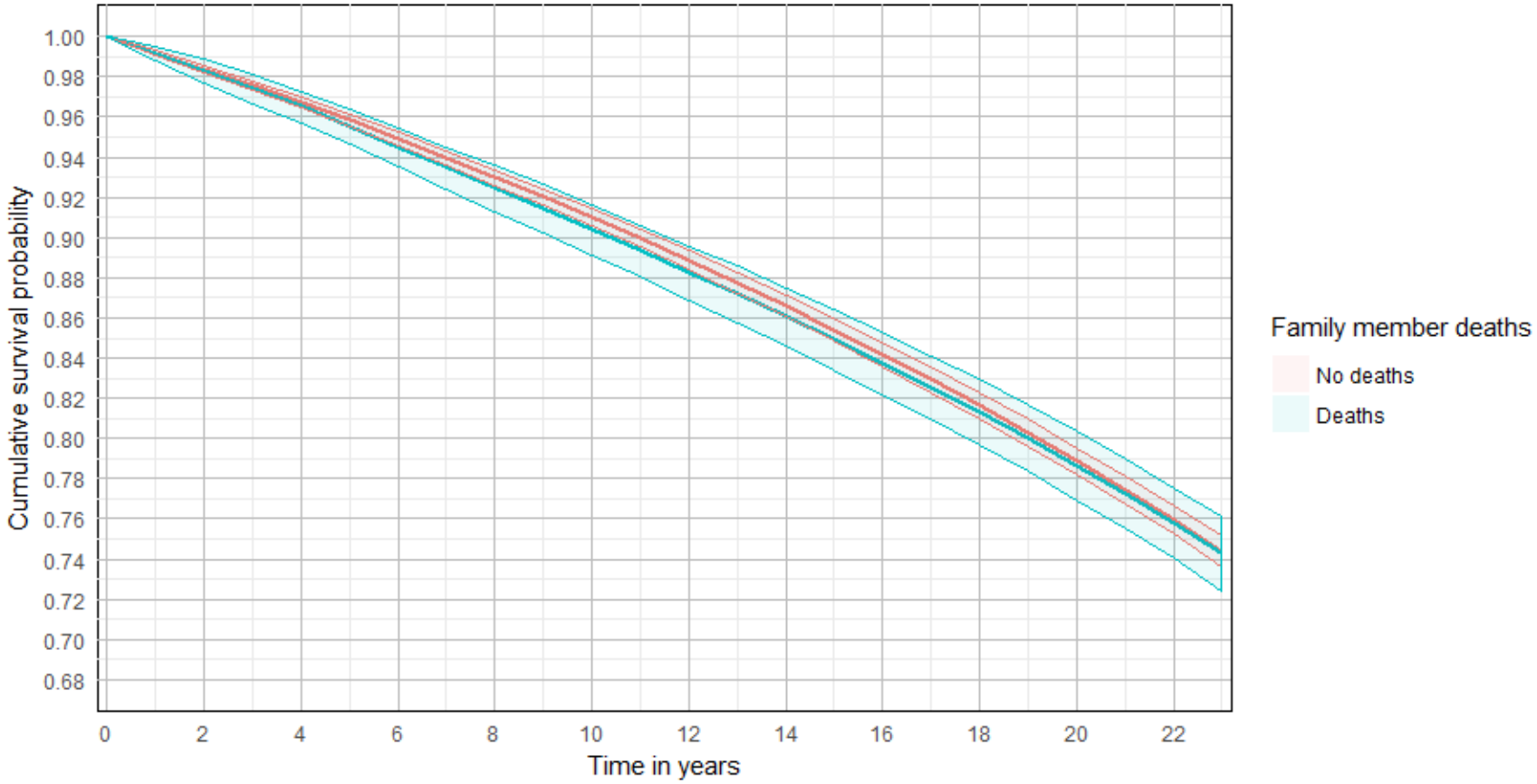

Figure 3.2 Cumulative survival probability among a cohort of the 1988 Spitak earthquake survivors during 1988-2012, by family member death status, using outcome data partially imputed by the Super Learner algorithm

[a] All the measures are standardized for age, age squared, gender, education in years, standard of living at the time of earthquake, place of residence (region), housing damage, and serious injury among the total baseline population

[b] Point estimates and 95% confidence limits are based on the 50th, 2.5th, and 97.5th percentile obtained from 5000 bootstrap sample estimates

Table 3.4 Marginal absolute effects of earthquake-related death of a family member on cumulative survival probability using outcome data partially imputed by the super learner algorithm among a cohort of the 1988 Spitak earthquake survivors

	Difference in cumulative survival probability [a] (95%CI) [b]	
	Earthquake-related death in the family	
Time (year)	**No death**	**Death of a family member**
1	Reference	0.000 (−0.004, 0.003)
2	Reference	−0.001 (−0.007, 0.005)
3	Reference	−0.001 (−0.009, 0.006)
5	Reference	−0.003 (−0.013, 0.006)
10	Reference	−0.006 (−0.019, 0.007)
15	Reference	−0.005 (−0.020, 0.010)
23	Reference	−0.001 (−0.020, 0.018)

[a] All the measures are standardized for age, age squared, gender, education in years, standard of living at the time of earthquake, place of residence (region), housing damage, and serious injury among the total baseline population

[b] Point estimates and 95% confidence limits are based on the 50th, 2.5th, and 97.5th percentile obtained from 5000 bootstrap sample estimates

Serious Injury

Plotting cumulative survival probabilities across those who sustained severe injury and those without serious injury showed no differences during the first 10 years of follow-up. From the second decade of the follow-up onwards, the two curves slowly diverged and the difference in cumulative survival probability of those with severe injury compared to those without such injury reached −0.033 (95%CI: −0.056, −0.009). Figure 3.3 provides the survival curves over time and Table 3.5 offers absolute effect estimates of injury on cumulative survival at selected time points.

Adjusting for the smoking and drinking habits of the study population at the baseline did not change the effect estimates and the underlying cumulative survival probabilities among the groups with and without serious injury (Table A-3.7).

In sensitivity analyses using outcome data partially imputed using random survival forest algorithm, we were able to replicate the results of the main analyses on the effect of serious injury on survival (Table A-3.8). Figure A-3.5 provides survival curves of the populations with and without serious injury. In analyses using data partially imputed by the score based multiple imputation algorithm, the results were somewhat different than those of the main analyses as well as of other sensitivity analyses and did not indicate a reduced survival probability among those with severe injuries in the later period of follow-up (Figure A-3.6 and Table A-3.9). Nevertheless, the 95% CIs of the survival estimates obtained from analyses of the data partially imputed by the score based multiple imputation algorithm were wide and indicated high levels of uncertainty.

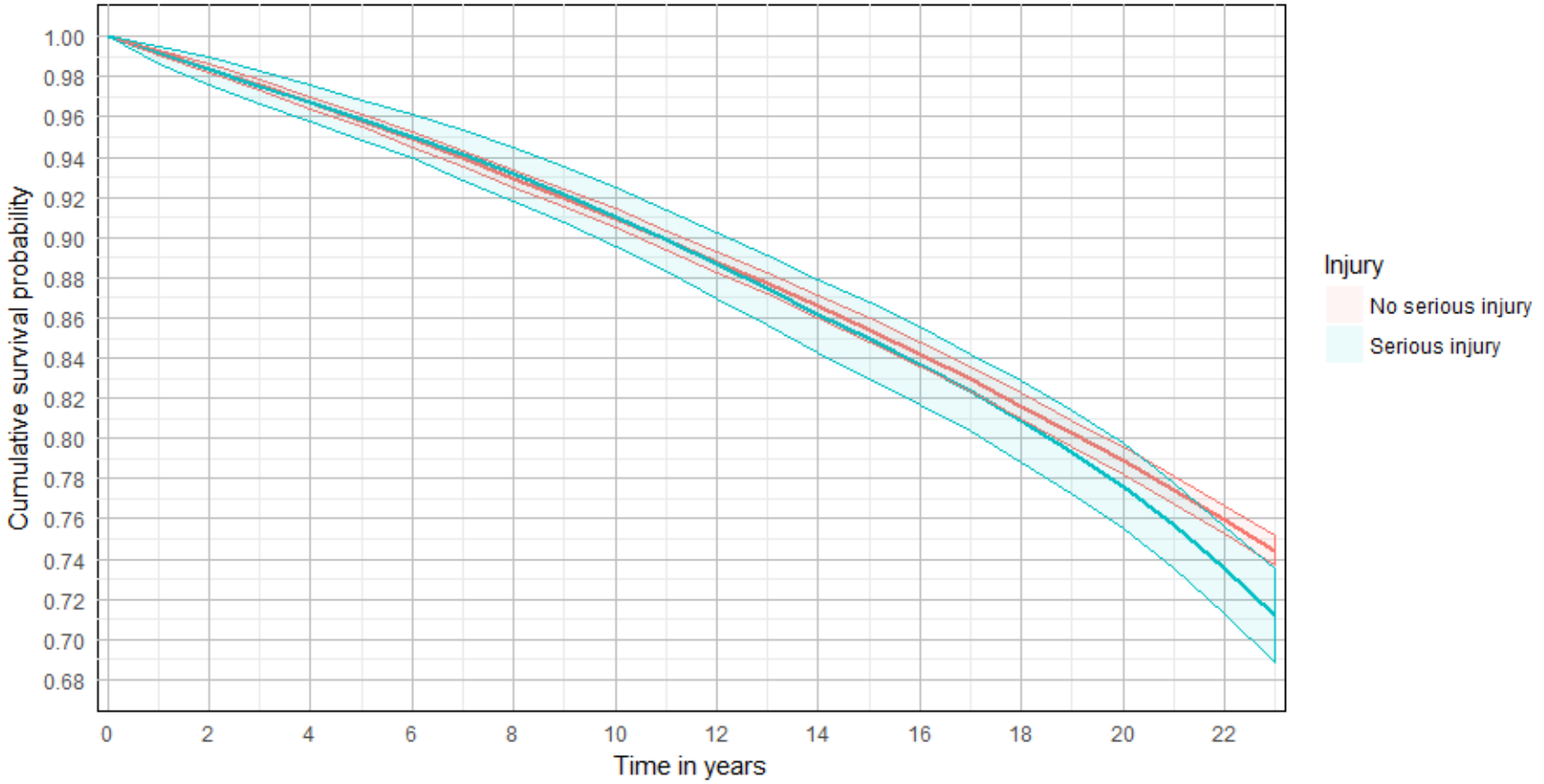

Figure 3.3 Cumulative survival probability among a cohort of the 1988 Spitak earthquake survivors during 1988-2012, by serious injury, using outcome data partially imputed by the Super Learner algorithm

[a] All the measures are standardized for age, age squared, gender, education in years, standard of living at the time of earthquake, place of residence (region), housing damage, and earthquake-related death of a family member among the total baseline population

[b] Point estimates and 95% confidence limits are based on the 50th, 2.5th, and 97.5th percentile obtained from 5000 bootstrap sample estimates

Table 3.5 Marginal absolute effects of earthquake caused serious injury on cumulative survival probability using outcome data partially imputed by the Super Learner algorithm among a cohort of the 1988 Spitak earthquake survivors

Difference in cumulative survival (95%CI) [a]	**Earthquake-related injury**	
Time (year)	**No serious injury**	**Serious injury**
1	Reference	−0.001 (−0.005, 0.003)
2	Reference	−0.001 (−0.008, 0.006)
3	Reference	0.000 (−0.009, 0.008)
5	Reference	0.001 (−0.010, 0.011)
10	Reference	0.001 (−0.015, 0.016)
15	Reference	−0.005 (−0.024, 0.014)
23	Reference	−0.033 (−0.056, −0.009)

[a] All the measures are standardized for age, age squared, gender, education in years, standard of living at the time of earthquake, place of residence (region), housing damage, and earthquake-related death of a family member among the total baseline population

[b] Point estimates and 95% confidence limits are based on the 50th, 2.5th, and 97.5th percentile obtained from 5000 bootstrap sample estimates

Effect decomposition of housing damage on cumulative survival, with respect to post-earthquake housing aid

Table 3.6 summarizes direct and indirect effects of earthquake-related housing damage on cumulative survival probability (i.e. CDE, RIE, MIE and PIE) using the four-way effect decomposition framework at selected time points. The CDE of housing damage with respect to post-earthquake housing aid on cumulative survival probability was consistently negative at different time points and was often larger in magnitude than the total effect of housing damage itself. The RIE and MIE of housing damage with respect to receiving permanent housing on cumulative survival probability fluctuated over time and were consistent with a mild positive impact. The PIE often was null, with data at a few time points suggestive of a possible mild

positive impact. The results for moderate housing damage and total housing destruction followed a similar pattern.

Adjusting for smoking and drinking habits of participants in the outcome and mediator models had no notable impact on estimated effects. Table A-3.10 provides results of mediation analyses adjusted for smoking and drinking habits. Table A-3.11 and Table A-3.12 present results from further sensitivity analyses based on outcome data partially imputed by the random survival forest algorithm and score based multiple imputation. At the end of the 23-year follow-up period, both set of sensitivity analyses replicated the pattern observed in the main analyses. The sensitivity analyses using outcome data partially imputed by the score based multiple imputation algorithm yielded estimates that were generally larger than the estimates obtained from the main analyses, as did the sensitivity analyses using outcome data partially imputed through the random survival forest algorithm. Nevertheless, the 95% confidence intervals across these analyses had notable overlaps.

Table 3.6 Marginal absolute effects of earthquake-related housing damage on cumulative survival probability using outcome data partially imputed by super learner among a cohort of the 1988 Spitak earthquake survivors

		Difference in cumulative survival probability [a] (95%CI) [b]		
		Housing damage		
Time (year)	**Effect measure**	**No damage**	**Moderate damage**	**Total destruction**
1	CDE	Reference	−0.003 (−0.005, −0.001)	−0.005 (−0.009, −0.002)
1	RIE	Reference	0.000 (−0.002, 0.003)	0.000 (−0.002, 0.003)
1	MIE	Reference	0.000 (−0.002, 0.002)	0.000 (−0.002, 0.002)
1	PIE	Reference	0.000 (−0.001, 0.001)	0.000 (−0.001, 0.001)
2	CDE	Reference	−0.005 (−0.009, −0.002)	−0.009 (−0.014, −0.004)
2	RIE	Reference	0.001 (−0.002, 0.004)	0.000 (−0.003, 0.004)
2	MIE	Reference	0.000 (−0.003, 0.003)	0.000 (−0.003, 0.004)
2	PIE	Reference	0.000 (−0.002, 0.002)	0.000 (−0.002, 0.002)
3	CDE	Reference	−0.007 (−0.012, −0.002)	−0.012 (−0.019, −0.006)
3	RIE	Reference	0.001 (−0.003, 0.004)	0.001 (−0.003, 0.005)
3	MIE	Reference	0.000 (−0.003, 0.004)	0.001 (−0.004, 0.004)
3	PIE	Reference	0.000 (−0.002, 0.003)	0.000 (−0.002, 0.003)
5	CDE	Reference	−0.009 (−0.014, −0.003)	−0.015 (−0.023, −0.008)
5	RIE	Reference	0.001 (−0.004, 0.006)	0.001 (−0.004, 0.005)
5	MIE	Reference	0.000 (−0.004, 0.004)	0.000 (−0.005, 0.004)
5	PIE	Reference	0.001 (−0.003, 0.004)	0.001 (−0.002, 0.004)
10	CDE	Reference	−0.010 (−0.017, −0.002	−0.019 (−0.030, −0.009)
10	RIE	Reference	0.001 (−0.005, 0.006)	0.000 (−0.007, 0.007)
10	MIE	Reference	−0.001 (−0.007, 0.006)	0.000 (−0.007, 0.006)
10	PIE	Reference	0.000 (−0.004, 0.004)	0.001 (−0.003, 0.006)
15	CDE	Reference	−0.011 (−0.021, −0.002)	−0.027 (−0.040, −0.014)
15	RIE	Reference	0.000 (−0.006, 0.007)	0.001 (−0.007, 0.009)
15	MIE	Reference	−0.001 (−0.008, 0.007)	0.000 (−0.007, 0.010)
15	PIE	Reference	0.001 (−0.005, 0.006)	0.000 (−0.005, 0.007)
23	CDE	Reference	−0.012 (−0.023, 0.000)	−0.040 (−0.055, −0.024)
23	RIE	Reference	0.000 (−0.007, 0.008)	0.004 (−0.006, 0.013)
23	MIE	Reference	−0.001 (−0.008, 0.007)	0.002 (−0.007, 0.013)
23	PIE	Reference	0.000 (−0.006, 0.006)	0.000 (−0.008, 0.007)

[a] All the measures are standardized for age, age squared, gender, education in years, standard of living at the time of earthquake, place of residence (region), earthquake-related death of a family member, and serious injury among the total baseline population

[b] Point estimates and 95% confidence limits are based on the 50th, 2.5th, and 97.5th percentile obtained from 5000 bootstrap sample estimates

3.4 Discussion

This study investigated the independent effect of housing damage, death of a family member, and severe injury on all-cause mortality among a population with a spectrum of exposure to the 1988 Spitak earthquake using hybrid data obtained from a prospective cohort study. The study revealed a link between housing damage or severe injury and all-cause mortality. In contrast, we did not find any difference in all-cause mortality rates among those who lost a family member to the earthquake and those who did not. Furthermore, this study decomposed the total effect of housing damage on all-cause mortality with respect to the mediating effect of receiving a permanent housing in the aftermath of the disaster and quantified direct and indirect effect measures.

The results show that those with total destruction of housing had substantially higher risk of all-cause mortality compared to those with no housing damage. A similar pattern with relatively smaller magnitude was observed when comparing those with moderate levels of housing damage compared to those without any housing damage. The effect of housing damage on mental health, and to a lesser extent on physical health outcomes have been previously studied (Armen K. Goenjian et al., 2018; Khachadourian et al., 2015). Mental health problems are often associated with increased risky behaviors, including smoking and alcohol use (Jacobson et al., 2008; Karestan C Koenen et al., 2006; McCauley, Killeen, Gros, Brady, & Back, 2012). Mental health conditions along with concurrent risky behaviors can directly or indirectly affect mortality rates. Hence, we believe that some of the observed increase in mortality rates is attributable to the effect of earthquake exposure mediated by preexisting mental and physical health conditions and their accompanying adverse health behaviors.

The main and sensitivity analyses were suggestive of no to possibly a very mild effect of the death of family member on all-cause mortality. The literature on the effect of death of family members on health outcomes is controversial. Adler (1943) in a 9-month follow up of 46 hospitalized victims of the Cocoanut Grove fire, found that those with a loss of family member or relative did not experience higher rates of mental health problems than their peers without personal losses. Similarly, Goenjian et al. (2018) in a study of the Spitak earthquake survivors, using data from the PEPSI cohort, found a positive association between death of a family member and PTSD severity score. Nevertheless, Murphy (1984) in a study of survivors of the Mount St. Helens disaster found no difference in physical health outcomes. In agreement with that, when assessing determinants of health-related quality of life 23 years after the Spitak earthquake using data from PEPSI cohort Khachadourian et al. (2015) found no association between death of a family member and health-related quality. Findings from the current study are in line with the disaster-related bereavement research in not finding an association between bereavement and physical health outcomes.

During the early period after the earthquake, the cumulative survival probability of those with severe injury did not differ from those without severe injury, however, during the period from 15 to 23 years after the earthquake, severe injury had a significant impact on mortality. By the end of the follow-up period, we observed substantially lower survival rates among those with severe injury compared to those without such injuries. While the injury directly and the traumatic experience (indirectly) have been suggested to be associated with mortality, absence of differential survival in the two groups might to a certain degree be due to selection bias (healthy survival effect) of those who had severe injury but did survive. Assuming injury was associated with destruction, and not necessarily with existing health conditions, then among those with a severe injury, individuals who survived would tend to be healthier and more resilient than those

who died because of the severe injury. Similarly severely injured survivors would thus be healthier than those who did not sustain a severe injury at all. Such selection bias (healthy survival effect) can result in underestimation of the potential negative effect of severe injury on mortality, resulting in a null effect, or in extreme cases, in a biased estimate across the null. Nonetheless, after a period of time, those healthy survivors might have a higher mortality rate, resulting in relatively higher mortality rates among those with severe injury during the later follow-up times. Although adjusting for smoking and drinking habits, as a proxy for overall health at the baseline, did not change the results, we cannot eliminate the possibility of the potential impact of the selection bias on the effect of severe injury on all-cause mortality. In light of these considerations, the estimates for the effect of severe injury on all-cause mortality should be interpreted with caution.

Mediation analysis decomposing the total effect of housing damage on all-cause mortality with respect to receiving permanent housing in the aftermath of the disaster, as expected, showed that a large portion of the observed effect of housing damage on all-cause mortality was not mediated through receiving permanent housing. The effect of interaction of receiving permanent housing with housing damage on all-cause mortality, captured through the mediated interaction and reference interaction effect measures, was the main source for the moderate positive impact of housing damage on all-cause mortality mediated through receiving permanent housing. Previous studies by Khachadourian et al. (2015) and Goenjian et al. (2018) studying health-related quality of life and PTSD among the Spitak earthquake survivors reported similar interactions between earthquake-related loss and post-earthquake support.

Although many of the study participants who suffered total housing destruction or moderate housing damage did not receive a permanent housing, when considering the higher likelihood of

receiving housing for those with total destruction of housing, as well as potential positive impact of such aid in the aftermath of the disaster, the result was in agreement with expectations.

The current study is one of the few investigating the effect of specific exposures on survival. The prospective design of the study and the availability of baseline data on a large number of participants with various levels of exposure to the earthquake along with the use of modern statistical techniques to create a hybrid dataset, allowed in-depth assessment of different exposure indicators of all-cause mortality. The large sample size allowed for more flexible and accurate estimation of survival curves, while the reporting of effect estimates in absolute values improved understandability as well as the practical implications of findings. Additionally, the flexibility of the models and the mediation analysis framework offered a platform for studying potential effect measure modification between exposures, mediators, and covariates while allowing identification of time dependent effect estimates (relaxing the proportional hazard assumption imposed in standard Cox proportional hazard models) and examine important patterns.

Several methodological limitations and shortcomings of the current study should be acknowledged. The use of hybrid data, namely simulation or imputation of outcome data, may have limited the accuracy of the results. Although the findings were robust under various sensitivity analyses using data simulated or imputed with alternative algorithms, all of these algorithms used the same underlying data to obtain the unobserved outcomes and could have inherited potential biases and methodological issues from the underlying data. All of the analyses assumed no unmeasured and uncontrolled confounding, an important, yet untestable, assumption. Assuming that destruction and the death of a family member were less likely to be related to individual characteristics, the assumption that no uncontrolled confounding factors for the effect of death of a family member and housing damage on all-cause mortality becomes

more realistic. Nevertheless, surviving severe injury could have been associated with baseline characteristics that were not fully controlled for in the analyses. Although the sensitivity analyses adjusting for smoking and alcohol use at the baseline were an attempt to address this concern, the possibility of residual confounding could not discard.

This study assessed the independent effect of selected earthquake exposures on all-cause mortality and quantified their effect over more than two decades of follow-up. The results shed light on the potential impact of these factors on all-cause mortality. Moreover, the mediation analyses showed one of the many possible pathways through which housing damage could potentially impact all-cause mortality among populations exposed to earthquake, expanding our understanding of the causal mechanism of housing damage on all-cause mortality. These findings can help public health practitioners and policy-makers to prioritize disaster preparedness efforts and improve the allocation of resources in the aftermath of disasters.

3.5 Appendix

Appendix figures

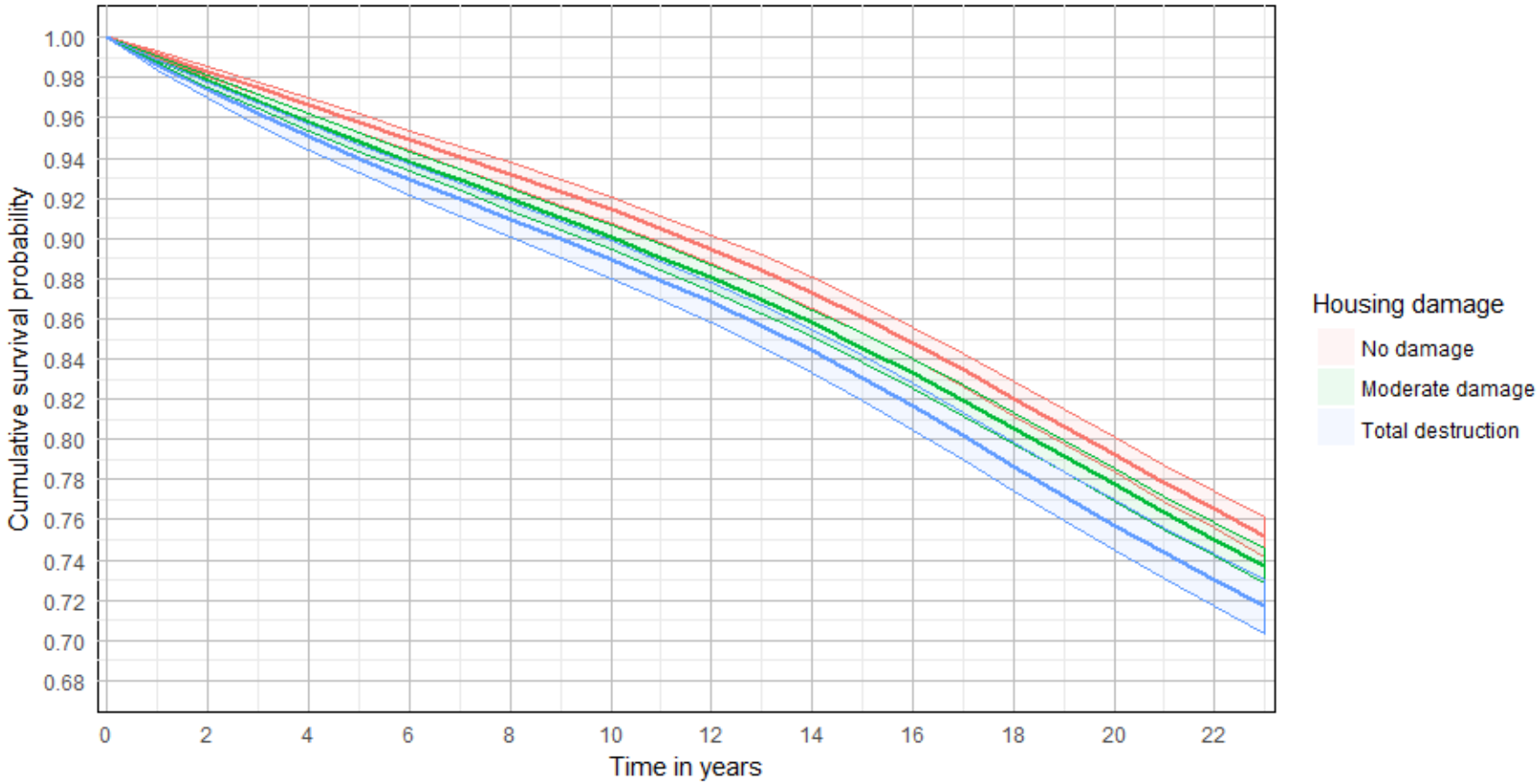

Figure A-3.1 Cumulative survival probability among a cohort of the 1988 Spitak earthquake survivors during 1988-2012, by housing damage levels, using outcome data partially imputed by random survival forest algorithm

[a] All the measures are standardized for age, age squared, gender, education in years, standard of living at the time of earthquake, place of residence (region), earthquake-related death of a family member, and serious injury among the total baseline population

[b] Point estimates and 95% confidence limits are based on the 50th, 2.5th, and 97.5th percentile obtained from 5000 bootstrap sample estimates

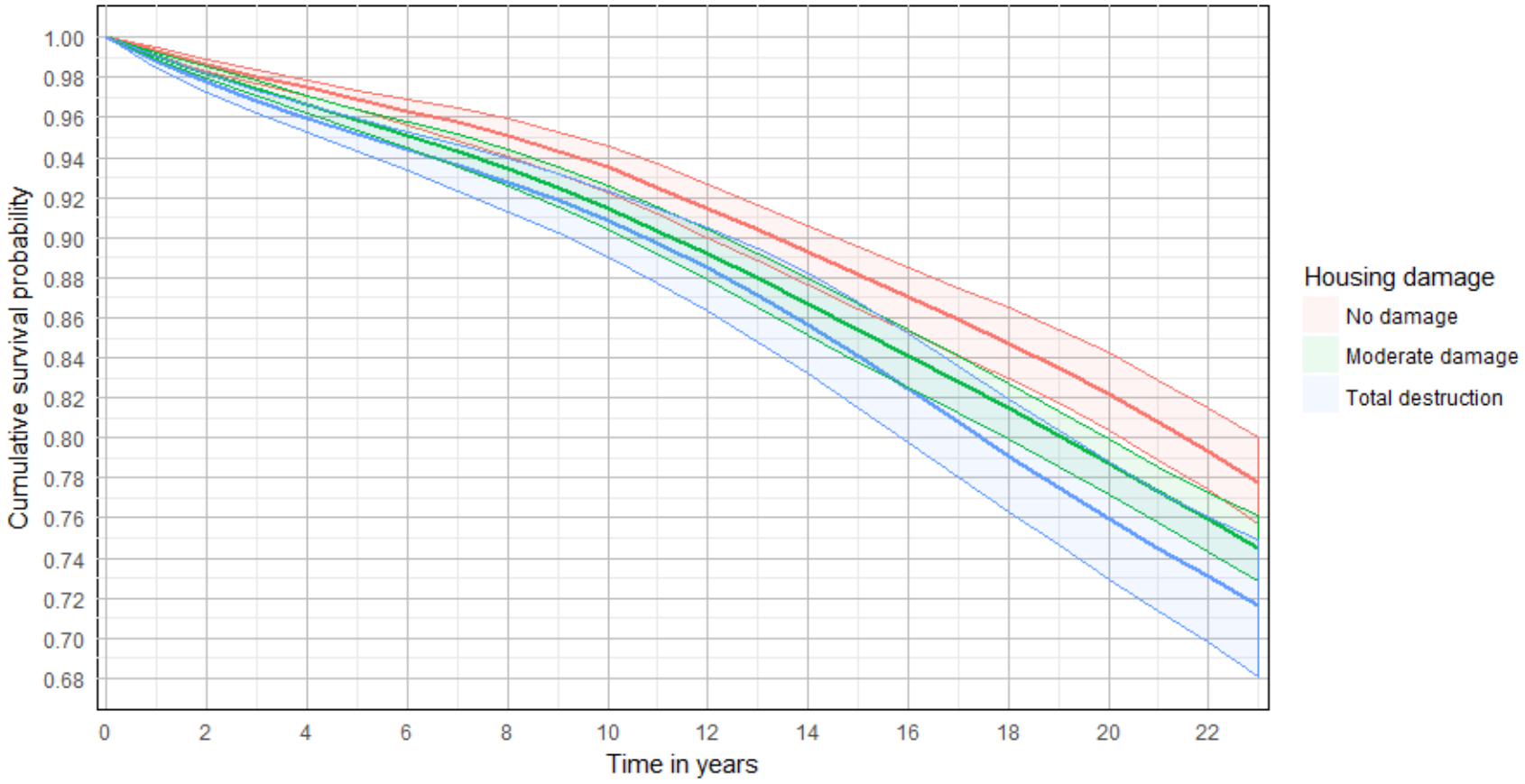

Figure A-3.2 Cumulative survival probability among a cohort of the 1988 Spitak earthquake survivors during 1988-2012, by housing damage levels, using outcome data partially imputed by risk score based multiple imputation algorithm

[a] All the measures are standardized for age, age squared, gender, education in years, standard of living at the time of earthquake, place of residence (region), earthquake-related death of a family member, and serious injury among the total baseline population

[b] Point estimates and 95% confidence limits are based on the 50th, 2.5th, and 97.5th percentile obtained from 5000 bootstrap sample estimates

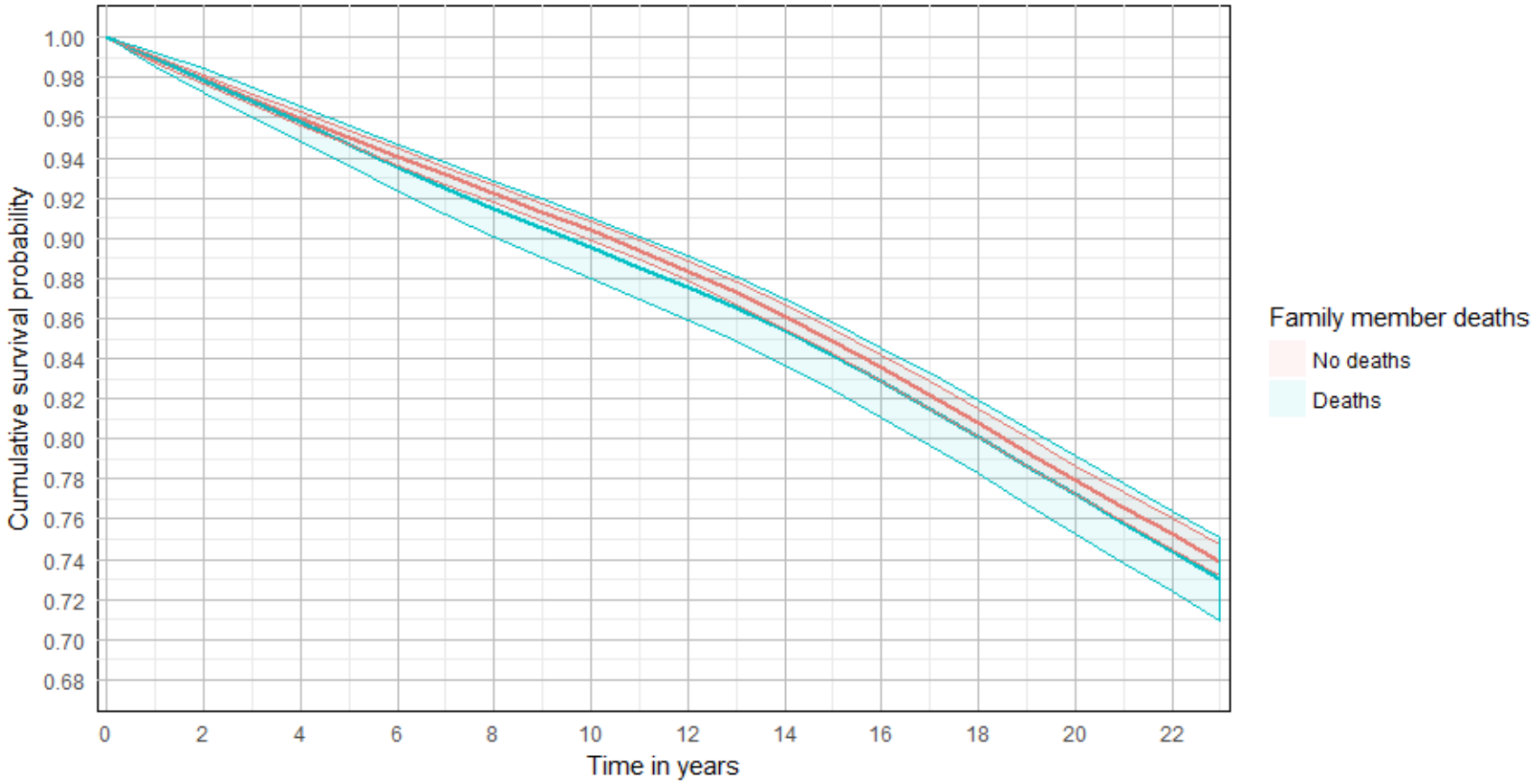

Figure A-3.3 Cumulative survival probability among a cohort of the 1988 Spitak earthquake survivors during 1988-2012, by family member death, using outcome data partially imputed by random survival forest algorithm

[a] All the measures are standardized for age, age squared, gender, education in years, standard of living at the time of earthquake, place of residence (region), housing damage, and serious injury among the total baseline population

[b] Point estimates and 95% confidence limits are based on the 50th, 2.5th, and 97.5th percentile obtained from 5000 bootstrap sample estimates

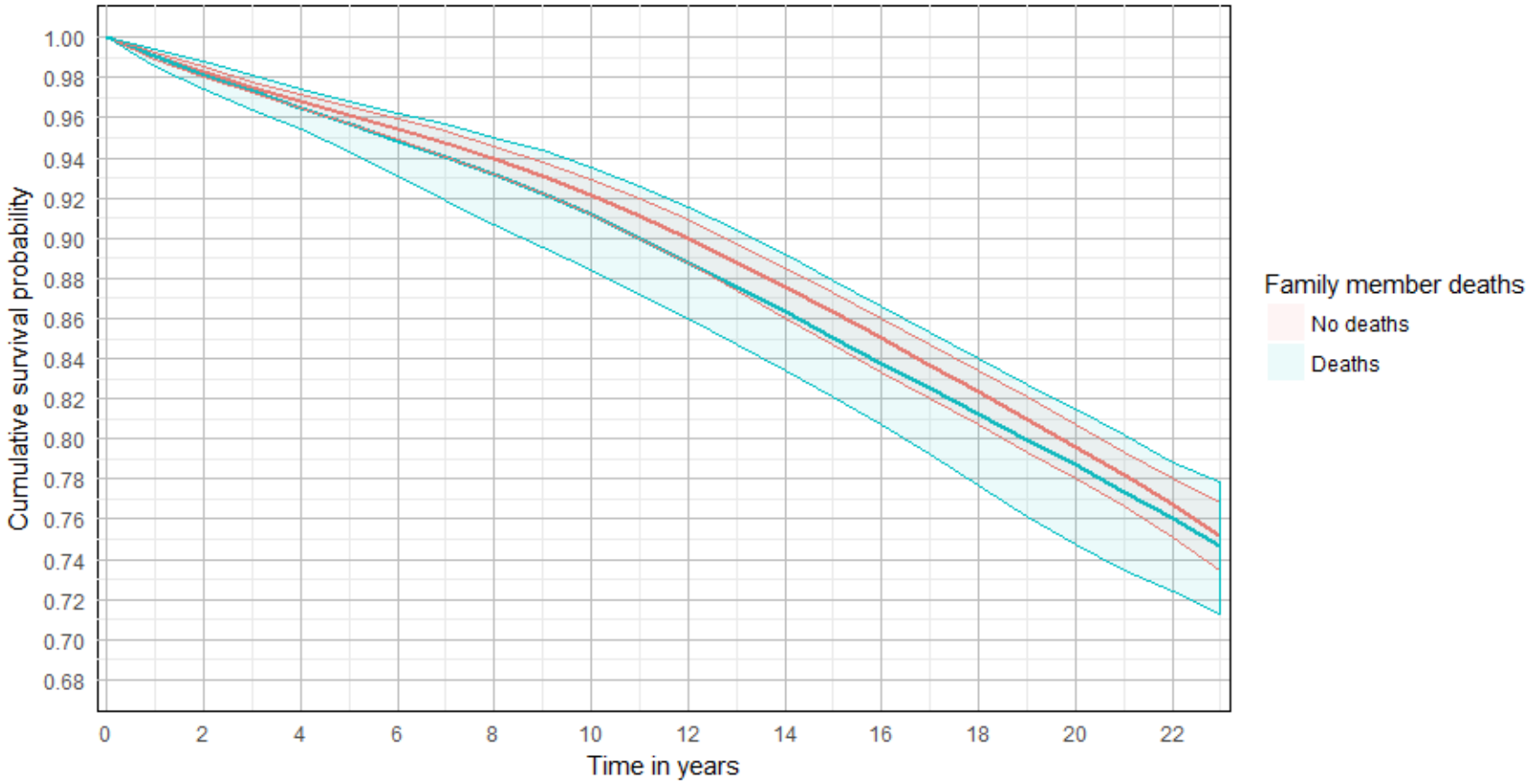

Figure A-3.4 Cumulative survival probability among a cohort of the 1988 Spitak earthquake survivors during 1988-2012, by family member death, using outcome data partially imputed by risk score based multiple imputation algorithm

[a] All the measures are standardized for age, age squared, gender, education in years, standard of living at the time of earthquake, place of residence (region), housing damage, and earthquake-related death of a family member among the total baseline population

[b] Point estimates and 95% confidence limits are based on the 50th, 2.5th, and 97.5th percentile obtained from 5000 bootstrap sample estimates

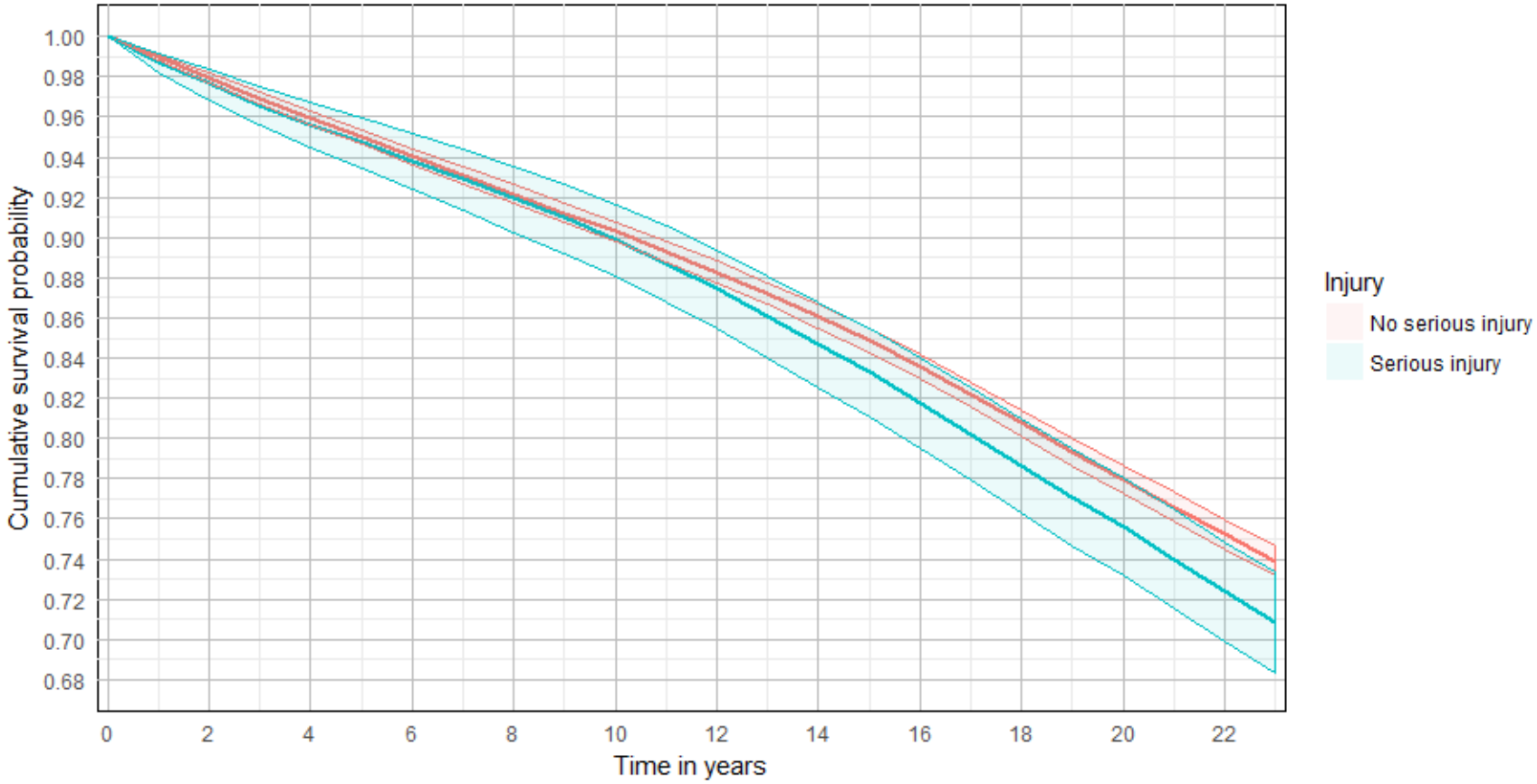

Figure A-3.5 Cumulative survival probability among a cohort of the 1988 Spitak earthquake survivors during 1988-2012, by injury status, using outcome data partially imputed by random survival forest algorithm

[a] All the measures are standardized for age, age squared, gender, education in years, standard of living at the time of earthquake, place of residence (region), housing damage, and earthquake-related death of a family member among the total baseline population

[b] Point estimates and 95% confidence limits are based on the 50th, 2.5th, and 97.5th percentile obtained from 5000 bootstrap sample estimates

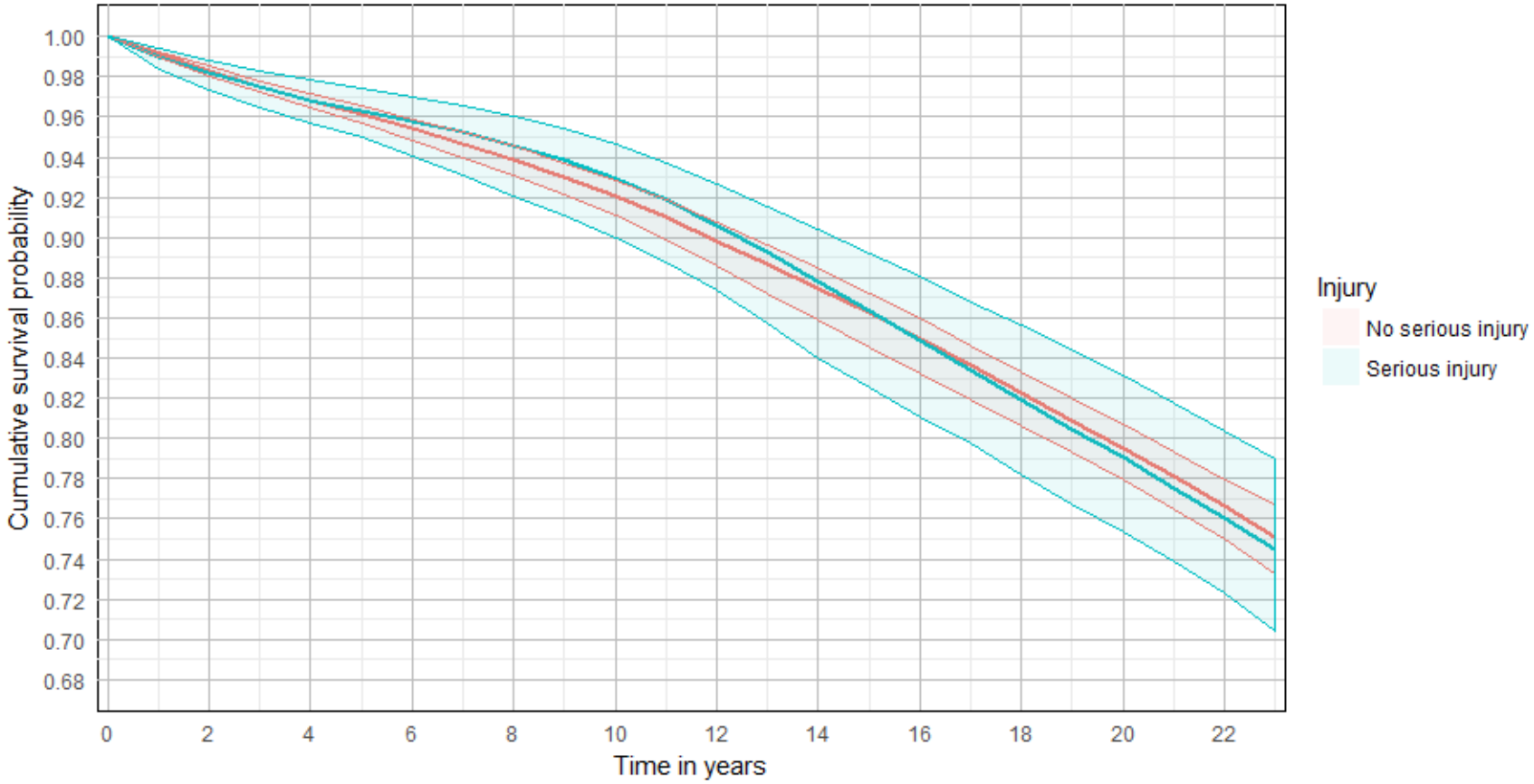

Figure A-3.6 Cumulative survival probability among a cohort of the 1988 Spitak earthquake survivors during 1988-2012, by injury status, using outcome data partially imputed by risk score based multiple imputation algorithm

[a] All the measures are standardized for age, age squared, gender, education in years, standard of living at the time of earthquake, place of residence (region), housing damage, and earthquake-related death of a family member among the total baseline population

[b] Point estimates and 95% confidence limits are based on the 50th, 2.5th, and 97.5th percentile obtained from 5000 bootstrap sample estimates

Appendix tables

Table A-3.1 Marginal absolute effects of earthquake-related housing damage on cumulative survival probability using outcome data partially imputed by the super learner algorithm among a cohort of the 1988 Spitak earthquake survivors

	Difference in cumulative survival probability [a] (95%CI) [b]		
	Housing damage		
Time (year)	**No damage**	**Moderate damage**	**Total destruction**
1	Reference	−0.003 (−0.005, 0.000)	−0.004 (−0.008, −0.001)
2	Reference	−0.005 (−0.008, −0.001)	−0.008 (−0.012, −0.003)
3	Reference	−0.006 (−0.010, −0.002)	−0.010 (−0.016, −0.005)
5	Reference	−0.008 (−0.013, −0.002)	−0.014 (−0.021, −0.007)
10	Reference	−0.009 (−0.016, −0.002)	−0.017 (−0.027, −0.008)
15	Reference	−0.011 (−0.019, −0.002)	−0.023 (−0.035, −0.012)
23	Reference	−0.010 (−0.021, 0.001)	−0.033 (−0.047, −0.019)

[a] All the measures are standardized for age, age squared, gender, education in years, standard of living at the time of earthquake, place of residence (region), earthquake-related death of a family member, serious injury, smoking, and drinking habit among the total baseline population

[b] Point estimates and 95% confidence limits are based on the 50th, 2.5th, and 97.5th percentile obtained from 5000 bootstrap sample estimates

Table A-3.2 Marginal absolute effects of earthquake-related housing damage on cumulative survival probability using outcome data partially imputed by random survival forest algorithm among a cohort of the 1988 Spitak earthquake survivors

	Difference in cumulative survival probability [a] (95%CI) [b]		
	Housing damage		
Time (year)	**No damage**	**Moderate damage**	**Total destruction**
1	Reference	−0.003 (−0.005, 0.000)	−0.005 (−0.008, −0.002)
2	Reference	−0.005 (−0.009, −0.001)	−0.009 (−0.014, −0.004)
3	Reference	−0.007 (−0.011, −0.002)	−0.013 (−0.019, −0.007)
5	Reference	−0.010 (−0.016, −0.003)	−0.018 (−0.026, −0.009)
10	Reference	−0.014 (−0.022, −0.005)	−0.025 (−0.036, −0.013)
15	Reference	−0.015 (−0.025, −0.005)	−0.030 (−0.044, −0.017)
23	Reference	−0.015 (−0.027, −0.002)	−0.035 (−0.051, −0.019)

[a] All the measures are standardized for age, age squared, gender, education in years, standard of living at the time of earthquake, place of residence (region), earthquake-related death of a family member, and serious injury among the total baseline population

[b] Point estimates and 95% confidence limits are based on the 50th, 2.5th, and 97.5th percentile obtained from 5000 bootstrap sample estimates

Table A-3.3 Marginal absolute effects of earthquake-related housing damage on cumulative survival probability using outcome data partially imputed by risk score based multiple imputation algorithm among a cohort of the 1988 Spitak earthquake survivors

	Difference in cumulative survival probability [a] (95%CI) [b]		
	Housing damage		
Time (year)	**No damage**	**Moderate damage**	**Total destruction**
1	Reference	−0.002 (−0.005, 0.001)	−0.005 (−0.009, −0.001)
2	Reference	−0.004 (−0.008, 0.000)	−0.009 (−0.014, −0.003)
3	Reference	−0.006 (−0.011, −0.001)	−0.012 (−0.019, −0.006)
5	Reference	−0.010 (−0.016, −0.003)	−0.017 (−0.027, −0.008)
10	Reference	−0.020 (−0.032, −0.006)	−0.027 (−0.042, −0.008)
15	Reference	−0.028 (−0.044, −0.010)	−0.042 (−0.064, −0.009)
23	Reference	−0.033 (−0.061, −0.007)	−0.062 (−0.089, −0.032)

[a] All the measures are standardized for age, age squared, gender, education in years, standard of living at the time of earthquake, place of residence (region), earthquake-related death of a family member, and serious injury among the total baseline population

[b] Point estimates and 95% confidence limits are based on the 50th, 2.5th, and 97.5th percentile obtained from 5000 bootstrap sample estimates

Table A-3.4 Marginal absolute effects of earthquake-related death of a family member on cumulative survival probability using outcome data partially imputed by super learner algorithm among a cohort of the 1988 Spitak earthquake survivors

	Difference in cumulative survival probability [a] (95%CI) [b]	
	Earthquake-related death in the family	
Time (year)	**No death**	**Death of a family member**
1	Reference	0.000 (−0.004, 0.003)
2	Reference	0.000 (−0.007, 0.005)
3	Reference	−0.001 (−0.009, 0.006)
5	Reference	−0.002 (−0.012, 0.006)
10	Reference	−0.005 (−0.018, 0.008)
15	Reference	−0.003 (−0.019, 0.012)
23	Reference	0.001 (−0.018, 0.019)

[a] All the measures are standardized for age, age squared, gender, education in years, standard of living at the time of earthquake, place of residence (region), earthquake-related death of a family member, serious injury, smoking, and drinking habit among the total baseline population

[b] Point estimates and 95% confidence limits are based on the 50th, 2.5th, and 97.5th percentile obtained from 5000 bootstrap sample estimates

Table A-3.5 Marginal absolute effects of earthquake-related death of a family member on cumulative survival probability using outcome data partially imputed by random survival forest algorithm among a cohort of the 1988 Spitak earthquake survivors

	Difference in cumulative survival probability [a] (95%CI) [b]	
	Earthquake-related death in the family	
Time (year)	**No death**	**Death of a family member**
1	Reference	0.000 (−0.004, 0.004)
2	Reference	0.000 (−0.007, 0.006)
3	Reference	−0.001 (−0.010, 0.007)
5	Reference	−0.004 (−0.015, 0.007)
10	Reference	−0.008 (−0.024, 0.007)
15	Reference	−0.007 (−0.025, 0.010)
23	Reference	−0.009 (−0.031, 0.012)

[a] All the measures are standardized for age, age squared, gender, education in years, standard of living at the time of earthquake, place of residence (region), housing damage, and serious injury among the total baseline population

[b] Point estimates and 95% confidence limits are based on the 50th, 2.5th, and 97.5th percentile obtained from 5000 bootstrap sample estimates

Table A-3.6 Marginal absolute effects of earthquake-related death of a family member on cumulative survival probability using outcome data partially imputed by risk score based multiple imputation algorithm among a cohort of the 1988 Spitak earthquake survivors

	Difference in cumulative survival probability [a] (95%CI) [b]	
	Earthquake-related death in the family	
Time (year)	**No death**	**Death of a family member**
1	Reference	−0.001 (−0.006, 0.003)
2	Reference	−0.002 (−0.009, 0.006)
3	Reference	−0.002 (−0.012, 0.006)
5	Reference	−0.005 (−0.019, 0.008)
10	Reference	−0.010 (−0.036, 0.015)
15	Reference	−0.012 (−0.040, 0.020)
23	Reference	−0.006 (−0.036, 0.023)

[a] All the measures are standardized for age, age squared, gender, education in years, standard of living at the time of earthquake, place of residence (region), housing damage, and serious injury among the total baseline population

[b] Point estimates and 95% confidence limits are based on the 50th, 2.5th, and 97.5th percentile obtained from 5000 bootstrap sample estimates

Table A-3.7 Marginal absolute effects of earthquake caused serious injury on cumulative survival probability using outcome data partially imputed by super learner algorithm among a cohort of the 1988 Spitak earthquake survivors

	Difference in cumulative survival probability [a] (95%CI) [b]	
	Earthquake-related injury	
Time (year)	**No serious injury**	**Serious injury**
1	Reference	−0.001 (−0.005, 0.003)
2	Reference	−0.001 (−0.008, 0.006)
3	Reference	0.000 (−0.009, 0.008)
5	Reference	0.001 (−0.010, 0.011)
10	Reference	0.001 (−0.015, 0.016)
15	Reference	−0.005 (−0.025, 0.014)
23	Reference	−0.033 (−0.057, −0.010)

[a] All the measures are standardized for age, age squared, gender, education in years, standard of living at the time of earthquake, place of residence (region), earthquake-related death of a family member, serious injury, smoking, and drinking habit among the total baseline population

[b] Point estimates and 95% confidence limits are based on the 50^{th}, 2.5^{th}, and 97.5^{th} percentile obtained from 5000 bootstrap sample estimates

Table A-3.8 Marginal absolute effects of earthquake caused serious injury on cumulative survival probability using outcome data partially imputed by random survival forest algorithm among a cohort of the 1988 Spitak earthquake survivors

	Difference in cumulative survival probability [a] (95%CI) [b]	
	Earthquake-related injury	
Time (year)	**No serious injury**	**Serious injury**
1	Reference	−0.002 (−0.007, 0.003)
2	Reference	−0.003 (−0.012, 0.005)
3	Reference	−0.003 (−0.014, 0.006)
5	Reference	−0.003 (−0.016, 0.010)
10	Reference	−0.004 (−0.023, 0.014)
15	Reference	−0.016 (−0.038, 0.007)
23	Reference	−0.031 (−0.057, −0.004)

[a] All the measures are standardized for age, age squared, gender, education in years, standard of living at the time of earthquake, place of residence (region), housing damage, and earthquake-related death of a family member among the total baseline population

[b] Point estimates and 95% confidence limits are based on the 50th, 2.5th, and 97.5th percentile obtained from 5000 bootstrap sample estimates

Table A-3.9 Marginal absolute effects of earthquake caused serious injury on cumulative survival probability using outcome data partially imputed by risk score based multiple imputation algorithm among a cohort of the 1988 Spitak earthquake survivors

	Difference in cumulative survival probability [a] (95%CI) [b]	
	Earthquake-related injury	
Time (year)	**No serious injury**	**Serious injury**
1	Reference	−0.001 (−0.007, 0.004)
2	Reference	−0.002 (−0.010, 0.006)
3	Reference	−0.001 (−0.011, 0.009)
5	Reference	0.002 (−0.011, 0.014)
10	Reference	0.008 (−0.016, 0.027)
15	Reference	0.002 (−0.040, 0.035)
23	Reference	−0.007 (−0.050, 0.044)

[a] All the measures are standardized for age, age squared, gender, education in years, standard of living at the time of earthquake, place of residence (region), housing damage, and earthquake-related death of a family member among the total baseline population

[b] Point estimates and 95% confidence limits are based on the 50th, 2.5th, and 97.5th percentile obtained from 5000 bootstrap sample estimates

Table A-3.10 Marginal absolute effects of earthquake-related housing damage on cumulative survival probability using outcome data partially imputed by random survival forest algorithm among a cohort of the 1988 Spitak earthquake survivors

		Difference in cumulative survival probability [a] (95%CI) [b]		
		Housing damage		
Time (year)	Effect measure	No damage	Moderate damage	Total destruction
1	CDE	Reference	−0.003 (−0.006, −0.001)	−0.005 (−0.008, −0.002)
1	RIE	Reference	0.000 (−0.002, 0.003)	0.000 (−0.002, 0.003)
1	MIE	Reference	0.000 (−0.002, 0.002)	0.000 (−0.002, 0.002)
1	PIE	Reference	0.000 (−0.001, 0.001)	0.000 (−0.001, 0.001)
2	CDE	Reference	−0.005 (−0.009, −0.001)	−0.009 (−0.014, −0.004)
2	RIE	Reference	0.001 (−0.002, 0.004)	0.000 (−0.003, 0.004)
2	MIE	Reference	0.000 (−0.003, 0.003)	0.000 (−0.003, 0.004)
2	PIE	Reference	0.000 (−0.002, 0.002)	0.000 (−0.002, 0.002)
3	CDE	Reference	−0.007 (−0.012, −0.002)	−0.012 (−0.018, −0.005)
3	RIE	Reference	0.001 (−0.003, 0.004)	0.001 (−0.004, 0.005)
3	MIE	Reference	0.000 (−0.003, 0.004)	0.001 (−0.004, 0.004)
3	PIE	Reference	0.000 (−0.002, 0.003)	0.000 (−0.002, 0.003)
5	CDE	Reference	−0.009 (−0.014, −0.003)	−0.015 (−0.022, −0.007)
5	RIE	Reference	0.001 (−0.004, 0.006)	0.000 (−0.005, 0.005)
5	MIE	Reference	0.000 (−0.004, 0.004)	0.000 (−0.005, 0.004)
5	PIE	Reference	0.001 (−0.002, 0.004)	0.001 (−0.002, 0.004)
10	CDE	Reference	−0.009 (−0.016, −0.001)	−0.018 (−0.028, −0.008)
10	RIE	Reference	0.000 (−0.005, 0.006)	0.000 (−0.007, 0.006)
10	MIE	Reference	−0.001 (−0.007, 0.005)	−0.001 (−0.007, 0.006)
10	PIE	Reference	0.000 (−0.003, 0.004)	0.001 (−0.003, 0.006)
15	CDE	Reference	−0.010 (−0.020, 0.000)	−0.024 (−0.037, −0.012)
15	RIE	Reference	0.000 (−0.006, 0.007)	0.000 (−0.008, 0.009)
15	MIE	Reference	−0.001 (−0.008, 0.007)	0.000 (−0.007, 0.010)
15	PIE	Reference	0.001 (−0.005, 0.006)	0.001 (−0.005, 0.007)
23	CDE	Reference	−0.010 (−0.021, 0.001)	−0.037 (−0.052, −0.022)
23	RIE	Reference	0.000 (−0.008, 0.007)	0.003 (−0.007, 0.013)
23	MIE	Reference	−0.001 (−0.008, 0.007)	0.001 (−0.008, 0.012)
23	PIE	Reference	0.000 (−0.006, 0.006)	0.000 (−0.007, 0.007)

[a] All the measures are standardized for age, age squared, gender, education in years, standard of living at the time of earthquake, place of residence (region), earthquake-related death of a family member, serious injury, smoking, and drinking habit among the total baseline population

[b] Point estimates and 95% confidence limits are based on the 50th, 2.5th, and 97.5th percentile obtained from 5000 bootstrap sample estimates

Table A-3.11 Marginal absolute effects of earthquake-related housing damage on cumulative survival probability using outcome data partially imputed by random survival forest algorithm among a cohort of the 1988 Spitak earthquake survivors

		Difference in cumulative survival probability [a] (95%CI) [b]		
		Housing damage		
Time (year)	**Effect measure**	**No damage**	**Moderate damage**	**Total destruction**
1	CDE	Reference	−0.003 (−0.006, 0.000)	−0.006 (−0.009, −0.002)
1	RIE	Reference	0.000 (−0.002, 0.003)	0.000 (−0.002, 0.004)
1	MIE	Reference	0.000 (−0.002, 0.003)	0.000 (−0.002, 0.002)
1	PIE	Reference	0.000 (−0.002, 0.002)	0.000 (−0.002, 0.002)
2	CDE	Reference	−0.005 (−0.009, −0.001)	−0.010 (−0.016, −0.005)
2	RIE	Reference	0.000 (−0.003, 0.004)	0.000 (−0.004, 0.004)
2	MIE	Reference	0.000 (−0.003, 0.004)	0.000 (−0.003, 0.004)
2	PIE	Reference	0.000 (−0.002, 0.003)	0.000 (−0.002, 0.003)
3	CDE	Reference	−0.007 (−0.013, −0.002)	−0.014 (−0.021, −0.007)
3	RIE	Reference	0.001 (−0.003, 0.005)	0.000 (−0.005, 0.005)
3	MIE	Reference	0.000 (−0.004, 0.004)	0.000 (−0.004, 0.005)
3	PIE	Reference	0.001 (−0.002, 0.004)	0.001 (−0.002, 0.004)
5	CDE	Reference	−0.011 (−0.017, −0.004)	−0.018 (−0.027, −0.010)
5	RIE	Reference	0.001 (−0.004, 0.006)	0.000 (−0.006, 0.005)
5	MIE	Reference	0.000 (−0.005, 0.005)	−0.001 (−0.006, 0.005)
5	PIE	Reference	0.001 (−0.003, 0.005)	0.001 (−0.002, 0.005)
10	CDE	Reference	−0.015 (−0.024, −0.005)	−0.024 (−0.037, −0.012)
10	RIE	Reference	0.001 (−0.005, 0.007)	−0.001 (−0.010, 0.006)
10	MIE	Reference	−0.001 (−0.007, 0.006)	−0.001 (−0.009, 0.005)
10	PIE	Reference	0.001 (−0.003, 0.005)	0.001 (−0.003, 0.007)
15	CDE	Reference	−0.017 (−0.027, −0.006)	−0.030 (−0.044, −0.015)
15	RIE	Reference	0.001 (−0.005, 0.008)	−0.001 (−0.010, 0.008)
15	MIE	Reference	−0.001 (−0.008, 0.007)	−0.001 (−0.009, 0.010)
15	PIE	Reference	0.001 (−0.005, 0.007)	0.001 (−0.005, 0.008)
23	CDE	Reference	−0.018 (−0.030, −0.005)	−0.040 (−0.056, −0.023)
23	RIE	Reference	0.002 (−0.006, 0.010)	0.003 (−0.008, 0.013)
23	MIE	Reference	0.000 (−0.008, 0.009)	0.001 (−0.009, 0.014)
23	PIE	Reference	0.001 (−0.006, 0.007)	0.002 (−0.007, 0.008)

[a] All the measures are standardized for age, age squared, gender, education in years, standard of living at the time of earthquake, place of residence (region), earthquake-related death of a family member, and serious injury among the total baseline population

[b] Point estimates and 95% confidence limits are based on the 50th, 2.5th, and 97.5th percentile obtained from 5000 bootstrap sample estimates

Table A-3.12 Marginal absolute effects of earthquake-related housing damage on cumulative survival probability using outcome data partially imputed by risk score based multiple imputation algorithm among a cohort of the 1988 Spitak earthquake survivors

		Difference in cumulative survival probability [a] (95%CI) [b]		
		Housing damage		
Time (year)	Effect measure	No damage	Moderate damage	Total destruction
1	CDE	Reference	−0.002 (−0.005, 0.001)	−0.005 (−0.009, −0.001)
1	RIE	Reference	0.000 (−0.002, 0.002)	0.000 (−0.002, 0.004)
1	MIE	Reference	0.000 (−0.002, 0.002)	0.000 (−0.002, 0.002)
1	PIE	Reference	0.000 (−0.002, 0.002)	0.000 (−0.001, 0.002)
2	CDE	Reference	−0.004 (−0.008, 0.000)	−0.009 (−0.015, −0.003)
2	RIE	Reference	0.000 (−0.003, 0.003)	0.000 (−0.004, 0.004)
2	MIE	Reference	0.000 (−0.003, 0.003)	0.000 (−0.003, 0.005)
2	PIE	Reference	0.000 (−0.002, 0.002)	0.000 (−0.002, 0.002)
3	CDE	Reference	−0.007 (−0.012, −0.001)	−0.013 (−0.020, −0.005)
3	RIE	Reference	0.001 (−0.003, 0.004)	0.001 (−0.004, 0.005)
3	MIE	Reference	0.000 (−0.003, 0.004)	0.001 (−0.004, 0.005)
3	PIE	Reference	0.000 (−0.002, 0.003)	0.000 (−0.002, 0.003)
5	CDE	Reference	−0.011 (−0.018, −0.004)	−0.019 (−0.029, −0.009)
5	RIE	Reference	0.001 (−0.003, 0.006)	0.001 (−0.005, 0.005)
5	MIE	Reference	0.000 (−0.004, 0.004)	0.000 (−0.005, 0.005)
5	PIE	Reference	0.001 (−0.002, 0.004)	0.001 (−0.002, 0.004)
10	CDE	Reference	−0.023 (−0.037, −0.008)	−0.029 (−0.046, −0.009)
10	RIE	Reference	0.002 (−0.004, 0.009)	0.001 (−0.007, 0.008)
10	MIE	Reference	0.000 (−0.006, 0.006)	0.000 (−0.008, 0.006)
10	PIE	Reference	0.001 (−0.003, 0.004)	0.001 (−0.003, 0.007)
15	CDE	Reference	−0.033 (−0.051, −0.014)	−0.045 (−0.071, −0.010)
15	RIE	Reference	0.004 (−0.004, 0.012)	0.001 (−0.009, 0.012)
15	MIE	Reference	0.000 (−0.006, 0.008)	0.001 (−0.008, 0.011)
15	PIE	Reference	0.001 (−0.005, 0.006)	0.001 (−0.005, 0.008)
23	CDE	Reference	−0.042 (−0.071, −0.016)	−0.072 (−0.100, −0.040)
23	RIE	Reference	0.006 (−0.005, 0.018)	0.007 (−0.008, 0.020)
23	MIE	Reference	0.002 (−0.006, 0.010)	0.004 (−0.007, 0.015)
23	PIE	Reference	0.000 (−0.006, 0.006)	0.001 (−0.008, 0.010)

[a] All the measures are standardized for age, age squared, gender, education in years, standard of living at the time of earthquake, place of residence (region), earthquake-related death of a family member, and serious injury among the total baseline population

[b] Point estimates and 95% confidence limits are based on the 50th, 2.5th, and 97.5th percentile obtained from 5000 bootstrap sample estimates

Chapter 4. Impact of earthquake exposure on and the mediating role of post-earthquake experiences in incident diabetes

Abstract

Studies directly investigating the impact of disaster exposure on specific physical health outcomes and diseases are scarce. Activation of the physiologic stress response is one of the potential pathways contributing to the risk of developing diabetes. Earthquakes are significant traumatic events which are often associated with substantial levels of stress and mental health problems. The current chapter examined the impact of earthquake-related exposure (i.e. housing damage, death of a family member due to the earthquake, and severe earthquake-related injury) on risk of diabetes using data from a prospective cohort of surviving adults with differential exposure levels to the 1988 earthquake in Armenia. A total of 1688 individuals were followed after the earthquake for up to 23 years. The discrete person-time data were used to conduct survival analysis adjusting for age, gender, education, standard of living, place of residence, earthquake related death in the family, and injury. Applying a Monte Carlo simulation, the estimated hazard of diabetes for individuals with different earthquake-related exposure levels were integrated over time, and were used to obtain the cumulative probability of remaining diabetes-free over the 23-year of follow-up period. Total destruction of housing was found to be associated with a lower cumulative probability of remaining diabetes-free over the 23-year of follow-up (−0.027 (95%CI: −0.103, 0.037)), while death of a family member and severe earthquake-related injury were not found to be associated with risk of diabetes. Decomposing the effect of housing damage on diabetes risk with respect to receiving permanent housing and job loss in the aftermath of the earthquake indicated that housing aid and job loss in the aftermath of the disaster mediate or mitigate a portion of the total effect of housing damage on diabetes. The findings show that total destruction of housing resulted in a

long-lasting increase in risk of diabetes. The results of this study can serve public health practitioners for identifying high-risk groups and targeting them with various interventions.

4.1 Introduction

The primary focus of post-disaster health research has been on mental and psychological health. Nevertheless, several studies have also examined associations between other health conditions and exposure to disasters (Armenian et al., 1998; Holen, 1991; Zaetta, Santonastaso, & Favaro, 2011). The immediate health effects of disaster may include injuries such as fractures, laceration, and disaster-specific conditions such as crush syndrome (Kantarci et al., 2002; Vanholder et al., 2007; Yzermans, van den Berg, & Dirkzwager, 2009).

Most physical health research among disaster survivors has focused on medically unexplained physical symptoms. This phenomenon is described as a group of physical symptoms that lead to care-seeking while a comprehensive health evaluation cannot attribute them to a specific clinical pathogenesis (Clauw et al., 2003; S. Wessely, Nimnuan, & Sharpe, 1999; Simon Wessely & White, 2004). Such symptoms are often found to be associated with common post-disaster mental health problems, including PTSD and depression (B van den Berg et al., 2008; Bellis van den Berg, Grievink, Yzermans, & Lebret, 2005).

In a study of a relatively small sample (n=100) of self-selected graduate students exposed to the San Francisco Bay Area earthquake of 1989, shortly after the event, 39% of participants reported palpitations (Cardena & Spiegel, 1993). Similarly, a study of a large sample of adult survivors of the 1999 Taiwan earthquake adult survivors whose houses were totally or partially destroyed by the earthquake, found a palpitation prevalence of 30% 2 years after the disaster (Chen et al., 2007). Despite the significant variations across study methodologies and findings,

the overall literature on the subject suggests an increase in the prevalence of physical symptoms among disaster survivors. In a study of workers exposed to the Hanshin-Awaji earthquake in 1995, Tainaka and colleagues (1998) estimated the prevalence of palpitation, poor appetite, fatigue, and other symptoms, immediately, at 3 months, and 18 months after the event, and found that those with severe damage to their housing had elevated stress-related symptoms, 18 months after the earthquake. More than 40 years after the Vajont dam disaster, Zaeta et. al (2011) observed significantly higher rates of gastroenterological, dermatological, respiratory and other adverse physical conditions among survivors than among the comparison group. In spite of a decline in the prevalence of the physical symptoms over time, other findings suggest that physical problems can persist for a long time (Yzermans et al., 2009). Such increased risk might be associated with long-term mental health conditions and could be directly attributable to the preceding disaster.

Less directly, several studies have explored the association of post-disaster psychopathology, particularly PTSD, with physical health outcomes and have demonstrated positive associations between them. For instance, Dirkzwager et al. (2007), studying a population affected by a firework depot explosion using electronic medical records and a self-rating PTSD scale, showed that participants with PTSD experienced more than a 90% increase in likelihood of developing a vascular problem. Similarly, in a 22-year prospective study of PTSD and incidence of type 2 diabetes in women, Roberts et al. (2015) found that PTSD symptoms were positively associated with incident diabetes, and that participants with the highest number of PTSD symptoms—compared to women who were not exposed to trauma—had almost a 2-fold higher risk of type 2 diabetes.

Studies directly investigating impact of disaster exposure on specific physical health outcomes and diseases are scarce. However, in a retrospective cohort study using a historic control, Aoki

et al. (2012) found an increased incidence of stroke, pneumonia, and cardiovascular disease, including heart failure, in the aftermath of the Great East Japan Earthquake of 2011.

Diabetes is one of the leading non-communicable physical morbidities associated with significant health and economic adversities. Activation of the physiologic stress response is suggested to be one of the pathways contributing to the risk of developing diabetes (Kelly & Ismail, 2015). Earthquake is a significant traumatic event which often has an adverse effect on the socioeconomic status of survivors and is associated with chronic stress and mental health problems (Armenian et al., 2000; Armen K. Goenjian et al., 2018). Thus, we hypothesized that earthquake-related exposure (housing damage, death of a family member, and serious injury) would increase diabetes risk. This study sought to test that hypothesis, and also to assess the role of receiving permanent housing and job loss as potential mediators of the effect of housing damage on incidence of diabetes among a cohort of survivors of the 1988 Spitak earthquake.

4.2 Methods

Study sample and variables

We used data from PEPSI study, a prospective follow-up study of more than 30,000 individuals with differential exposure levels to the 1988 Spitak earthquake. The analyses were limited to 1710 participants who were intended to be followed at the Phase IV of the study, in 2012. Of the total participants, 1688 did not report diagnosis of diabetes at the time of the earthquake, hence were considered at risk and eligible for current study's purposes. Further description of the study design and methods is provided in Chapter 2 and elsewhere (Khachadourian et al., 2016).

Outcome

The physician diagnosed self-reported diabetes data were collected throughout the follow-up phases. Year of diagnosis was recorded and verified through subsequent phases of follow-up.

Exposures

Housing damage variable was a categorical variable defined as a) no damage, b) minor damages/non-structural damages, and c) extensive damage/structural damages/total destruction. Number of nuclear family deaths due to earthquake was collected and since there were only a few individuals with multiple deaths in the family, the variable was dichotomized to no death vs. any death. We created a binary variable (yes/no) to summarize earthquake-related serious injury to the study participant. Scratches and minor cuts were not considered as a serious injury.

Mediator (Post-earthquake variables)

Receiving permanent housing during the three years following the earthquake (measured during 1989–1991) was considered as a potential mediator. Job loss after the earthquake was defined as an event of job loss after which, despite seeking employment, no new job was found for more than six months.

Covariates

Data on age, gender, education, employment, pre-earthquake residence location and type, and living standard were collected at Phase I and were verified during the follow-up phases. The phase I questionnaire also collected data about pre-earthquake family structure.

Age was a continuous variable recorded in years. Education was recorded as a continuous variable reflecting the number of years of formal education. The pre-earthquake employment

variables included a binary variable indicating whether the participant was employed or not and a categorical variable pertaining job of those employed. Total number of household member (continuous variable), and number of children living in the household (a continuous variable) were variables summarizing family structure of participants. Data on pre-earthquake life style measures including consumption of alcohol, smoking habits, and physical activity were also available and were considered in the sensitivity analyses.

Statistical analysis

We used frequencies, proportion, means and standard deviations to summarize the baseline characteristics of the study population by earthquake-related exposure variables. The analytical sample was weighted to represent those individuals who were adults at the time of the earthquake (18 and above) but were censored during the first three years of the follow-up.

We selected 10,000 bootstrap sample from the original person-oriented dataset and restructured all the bootstrap samples into person-period format (Singer & Willett, 1993). Assuming consistency, uncontrolled confounding (conditional exchangeability), and positivity, we fit a pooled logistic regression to assess the independent impact of earthquake-related housing damage, earthquake-related injury, and earthquake-related death in the nuclear family on risk of diabetes. To estimate the time-dependent intercepts, we used natural cubic splines of time since the earthquake [in years], with five knots (at years 2, 6, 12, 17, and 22) to create a smoothing function. Although the discrete time-to-event analysis can easily accommodate time varying effect estimates and relax the proportional hazard assumption inherited to standard cox proportional hazard regression models, after assessing the proportional hazard assumption and considering the limited sample size we decided to eliminate interaction between the covariates and baseline hazards (time variables). The final model was adjusted for potential confounders of

the earthquake-related exposures' effects on diabetes, including age, age squared, gender, education in years, standard of living at the time of earthquake, and place of residence (region).

The roles of post-earthquake provision of permanent housing and job loss as potential mediator of the effect of housing damage on risk of diabetes were evaluated within a mediation analysis framework. Structures with multiple mediators do not readily fit within the natural decomposition framework. The natural direct and indirect effects, similar to the four-way decomposition described and applied in Chapters 2 and 3, require the cross-world assumption. Those models do not also allow for exposure-induced mediator-outcome confounding. Therefore, we employed an interventional approach for path-specific effect decomposition which relaxes the assumption of no exposure-induced mediator–outcome confounding and yields direct and indirect effect estimates summing to the total effect (Vansteelandt & Daniel, 2017).

The analyses were conducted in SAS 9.4. All effect estimates were based on the cumulative survival probabilities obtained by integration of the hazards over time via Monte Carlo simulations. The final effect estimates on the absolute scale (i.e. survival difference) were reported for year 1, 2, 3, 5, 10, and 23 of the follow-up. The point estimates of the measures of interest and their 95% confidence intervals were based on the corresponding percentile of ordered estimates from the 10,000 bootstrap samples.

Sensitivity analysis

The robustness of the results was evaluated via sensitivity analyses, further adjusting for individuals' smoker/non-smoker status and body mass index (BMI) at the time of the earthquake.

4.3 Results

Study participants had an average age of 38.8 years, with as average of 11.7 years of formal education, and were 57.8% female. Almost two-thirds of participants had an average standard of living at the time of the earthquake, one-third reported below-average, and only 3.9% reported above-average standard of living. Slightly less than one-third of participants were smokers, and s similar proportion of the sample reported drinking alcohol. On average, one in six reported regular exercise at the time of the earthquake. More than 90% of the individuals in the study resided in urban areas at the time of the earthquake. Gyumri and Vanadzor residents accounted for 40.2% and 40.1% of the sample, while Spitak and other areas constituted 16.5% and 3.1% of the study population, respectively. Table 4.1 provides further details on characteristics of the study participants at the time of the earthquake.

Participants living in Spitak, the city closest to the epicenter of the earthquake had the highest levels of exposure and losses. Three out of four (73.8%) participants in Spitak had their housing totally destroyed due to the earthquake, while the earthquake completely destroyed the housing of 21.6% and 5.5% of participants living in Gyumri and Vanadzor, respectively. About one in ten participants (9.7%) lost a family member to the earthquake: 32.6% in Spitak, 9.3% in Gyumri, and 0.6% in Vanadzor. Earthquake-related severe injury was the highest in Spitak (11.5%), followed by Gyumri (3.8%), and Vanadzor (2.4%).

During the period from earthquake until the end of 1991, more than a quarter of participants lost their job and despite their willingness to work, they remained unemployed for 6 months or more. In 1989–1991, permanent housing was provided to 21.2% of participants from Spitak, 16.4% from Gyumri, and 7.5% from Vanadzor. Further details about earthquake exposure and post-earthquake experience in the study sample by place of residence at the time of the earthquake are provided in Table 4.2.

Table 4.1 Baseline characteristics (at the time of the earthquake, in 1988) of the sample of eligible participants attempted to be followed in 2012

Characteristics	Total sample (N=1688)
Age, years mean (SD)	38.8 (14.3)
Gender, n (%)	
Male	712 (42.2)
Female	976 (57.8)
Education, years mean (SD)	11.7 (2.9)
Living standard in 1988, n (%)	
Above average	66 (3.9)
Average	1078 (63.9)
Below average	544 (32.2)
Smoking status in 1988, n (%)	
Smoker	496 (29.4)
Non-smoker	1192 (70.6)
Drinking in 1988, n (%)	
Yes	493 (29.2)
No	1195 (70.8)
Regular exercise, n (%)	
Yes	285 (16.9)
No	1403 (83.1)
Area of residence, n (%)	
Spitak	279 (16.5)
Gyumri	679 (40.2)
Vanadzor	677 (40.1)
Other areas	53 (3.1)
Type of residence, n (%)	
Rural	124 (7.4)
Urban	1564 (92.6)

Table 4.2 Participants baseline characteristics by region of residence at the time of the earthquake, in 1988

Earthquake-related exposure and post-earthquake events	Spitak	Gyumri	Vanadzor	Total
Housing damage, n (%)				
No damage	14 (5.0)	305 (41.7.)	252 (37.2)	571 (33.8)
Moderate damage	59 (21.2)	269 (36.7)	388 (57.3)	716 (42.2)
Total destruction	206 (73.8)	158 (21.6)	37 (5.5)	401 (23.8)
Earthquake-related death in the nuclear family, n (%)				
Yes	91 (32.6)	68 (9.3)	4 (0.6)	163 (9.7)
No	188 (67.4)	664 (90.7)	673 (99.4)	1525 (90.3)
Earthquake caused serious injury, n (%)				
Yes	32 (11.5)	28 (3.8)	16 (2.4)	76 (4.5)
No	247 (88.5)	704 (96.2)	661 (97.6)	1612 (95.5)
Job loss for more than 6 month during 1989-1991, n (%)				
Yes	116 (41.6)	176 (24.0)	141 (10.8)	433 (25.7)
No	163 (58.4)	556 (76.0)	79 (79.2)	1255 (74.3)
Receiving a housing as an aid during 1989-1991, n (%)				
Yes	59 (21.2)	120 (16.4)	51 (7.5)	230 (13.6)
No	220 (78.8)	612 (83.6)	626 (92.5)	1458 (86.4)

Effect of earthquake-related exposures on incident diabetes

Housing damage

In the initial analysis standardized for the covariates in the total population, the 23-year cumulative probability of remaining diabetes-free among those with no or moderate housing damage, respectively, was 0.922 (95%CI: 0.885, 0.952) and 0.932 (95%CI: 0.901, 0.957), while those with total destruction of their housing had a cumulative diabetes-free probability of 0.895 (95%CI: 0.824, 0.946). Figure 3.1 presents the cumulative diabetes-free probability of the study population over the follow-up period by level of housing damage.

Those who experienced moderate housing damage had a relatively similar cumulative probability of remaining diabetes-free compared to the population with no damage to their housing. The difference in cumulative survival probability between these two groups, 23 years after the earthquake, was 0.009 (95%CI: −0.031, 0.054). The difference in the cumulative probability of remaining diabetes free in the population with a total housing destruction compared to the population with no housing damage at the end of the 23-year follow up was −0.027 (95%CI: −0.103, 0.037). Table 4.3 provides the marginal absolute effect estimates of housing damage on the cumulative probability of remaining diabetes-free at selected time points.

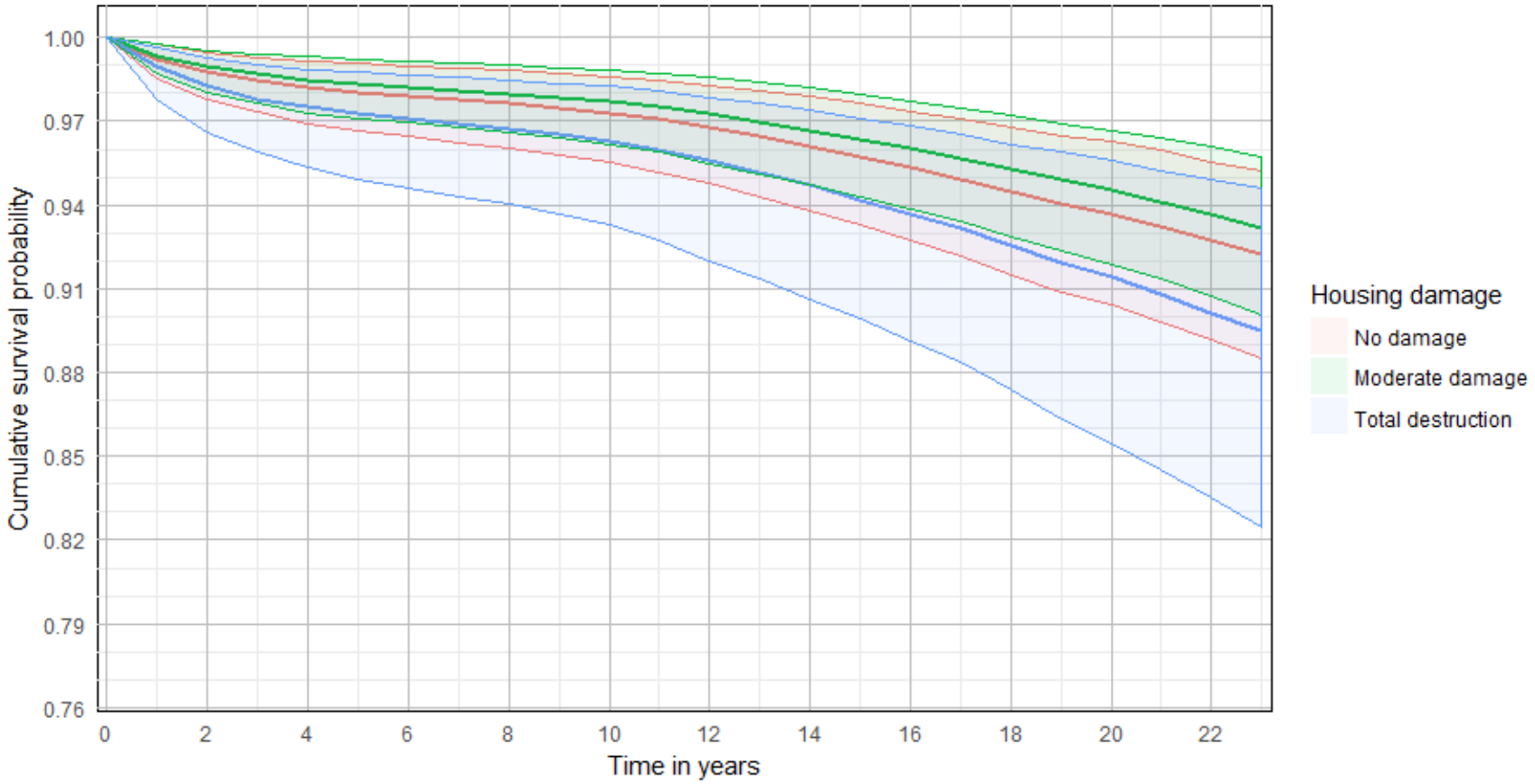

Figure 4.1 Cumulative survival probability from risk of diabetes among a cohort of the 1988 Spitak earthquake survivors during 1988-2012, by housing damage

[a] All the measures are standardized for age, age squared, gender, education in years, standard of living at the time of earthquake, place of residence (region), earthquake-related death of a family member, and serious injury among the total baseline population

[b] Point estimates and 95% confidence limits are based on the 50th, 2.5th, and 97.5th percentile obtained from 10,000 bootstrap sample estimates

Table 4.3 Marginal absolute effects of earthquake-related housing damage on cumulative survival probability from risk of diabetes among a cohort of the 1988 Spitak earthquake survivors

	Difference in cumulative survival probability [a] (95%CI) [b]		
	Housing damage		
Time (year)	**No damage**	**Moderate damage**	**Total destruction**
1	Reference	0.001 (−0.006, 0.008)	−0.003 (−0.014, 0.006)
2	Reference	0.002 (−0.008, 0.012)	−0.005 (−0.021, 0.008)
3	Reference	0.002 (−0.009, 0.014)	−0.006 (−0.024, 0.010)
5	Reference	0.003 (−0.011, 0.018)	−0.008 (−0.033, 0.013)
10	Reference	0.004 (−0.014, 0.023)	−0.010 (−0.042, 0.016)
15	Reference	0.006 (−0.019, 0.034)	−0.015 (−0.061, 0.024)
23	Reference	0.009 (−0.031, 0.054)	−0.027 (−0.103, 0.037)

[a] All the measures are standardized for age, age squared, gender, education in years, standard of living at the time of earthquake, place of residence (region), earthquake-related death of a family member, and serious injury among the total baseline population

[b] Point estimates and 95% confidence limits are based on the 50th, 2.5th, and 97.5th percentile obtained from 10,000 bootstrap sample estimates

Sensitivity analyses further adjusting for smoking and BMI of participants at the baseline yielded results consistent with those presented above. Figure A-4.1 and Table A-4.1 in the appendix presents the cumulative survival curves and marginal effect estimates from the sensitivity analyses.

Death of a family member

The cumulative probability of remaining diabetes-free among those who lost a family member to the earthquake and among those who did not experience such a loss did not differ appreciably over the follow-up time (0.919 vs. 0.920), (Figure 4.2). Table 4.4 presents the marginal absolute effect estimates of loss of a family member to the earthquake on cumulative probability of remaining diabetes free at selected time points over the follow-up period.

We did not observe any significant change in the effect estimates of death of a family member on cumulative probability of remaining diabetes-free over time in sensitivity analyses adjusting for smoking and BMI of study participants at the baseline. The cumulative probability of remaining diabetes-free plotted along with effect estimates from these analyses are presented in Figures A-4.2 and Table A-4.2, respectively.

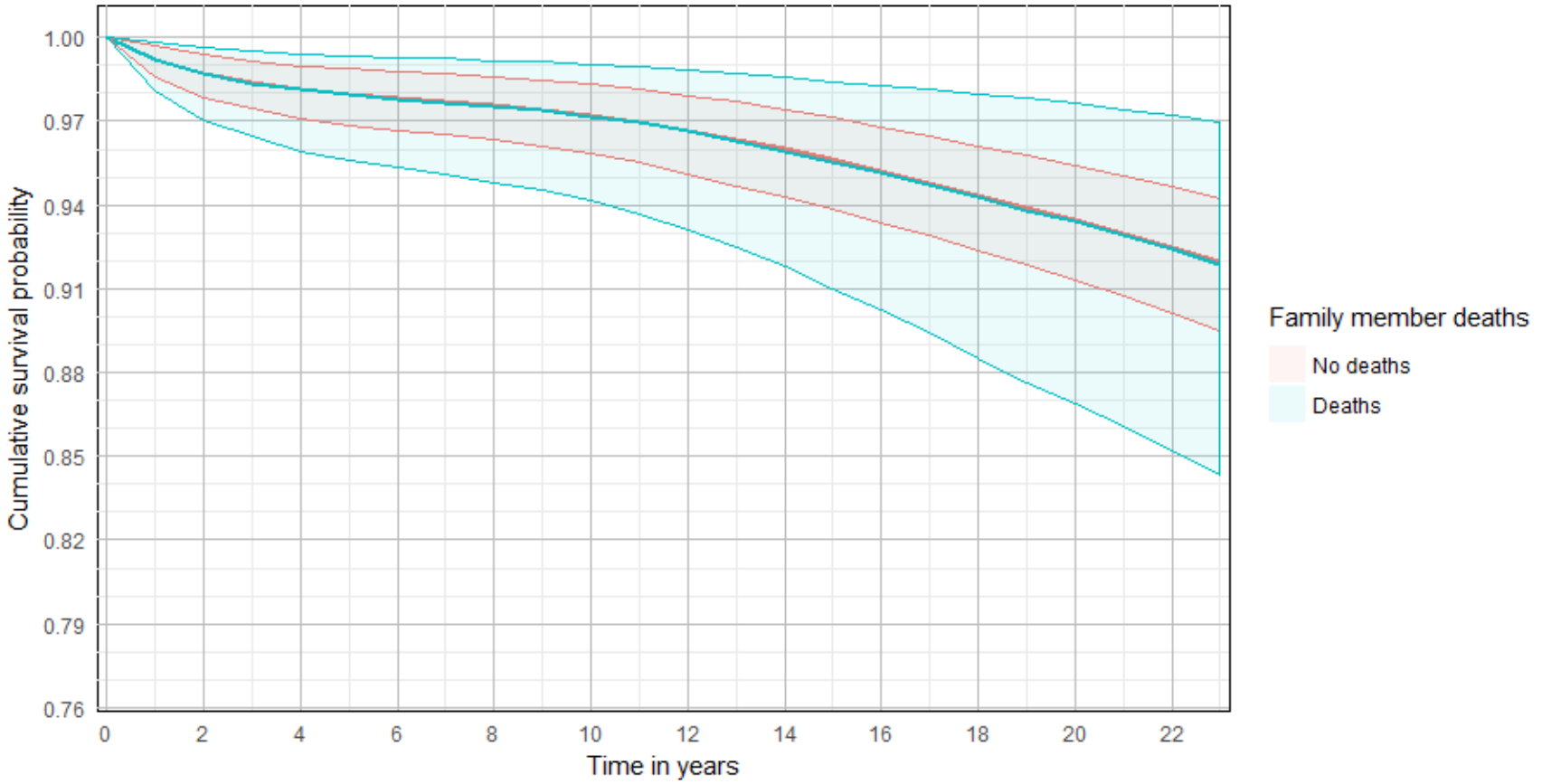

Figure 4.2 Cumulative survival probability from risk of diabetes among a cohort of the 1988 Spitak earthquake survivors during 1988-2012, by family member death status

[a] All the measures are standardized for age, age squared, gender, education in years, standard of living at the time of earthquake, place of residence (region), housing damage, and serious injury among the total baseline population

[b] Point estimates and 95% confidence limits are based on the 50th, 2.5th, and 97.5th percentile obtained from 10,000 bootstrap sample estimates

Table 4.4 Marginal absolute effects of earthquake-related death of a family member on cumulative survival probability rom risk of diabetes among a cohort of the 1988 Spitak earthquake survivors

	Difference in cumulative survival probability [a] (95%CI) [b]	
	Earthquake-related death in the family	
Time (year)	**No death**	**Death of a family member**
1	Reference	0.000 (−0.011, 0.008)
2	Reference	0.000 (−0.016, 0.011)
3	Reference	0.000 (−0.019, 0.014)
5	Reference	−0.001 (−0.024, 0.017)
10	Reference	−0.001 (−0.032, 0.022)
15	Reference	−0.001 (−0.048, 0.033)
23	Reference	−0.001 (−0.079, 0.055)

[a] All the measures are standardized for age, age squared, gender, education in years, standard of living at the time of earthquake, place of residence (region), housing damage, and serious injury among the total baseline population

[b] Point estimates and 95% confidence limits are based on the 50th, 2.5th, and 97.5th percentile obtained from 10,000 bootstrap sample estimates

Serious Injury

The cumulative probability of remaining diabetes-free curves among the population who sustained severe injury and those who did not experience serious injury suggested some differences. The cumulative probabilities of remaining diabetes-free at the end of the 23-year follow-up of 0.943 (95%CI: 0.865, 1.000) among those with severe injury and 0.919 (95%CI: 0.894, 0.941) among those with no serious injury yielded a difference in cumulative probability of remaining diabetes free of 0.024 (95%CI: −0.056, 0.084).

The sensitivity analyses adjusting for smoking and drinking of the study participants at the baseline did slightly decreased the gap in the cumulative probability of remaining diabetes-free between those with and without serious injury (0.018 (95%CI: 0.070, 0.083)). Figure A-4.3

provides survival curves of these populations and table A-4.3 lists the corresponding effect estimates calculated by contrasting the survival curves obtained from the sensitivity analyses.

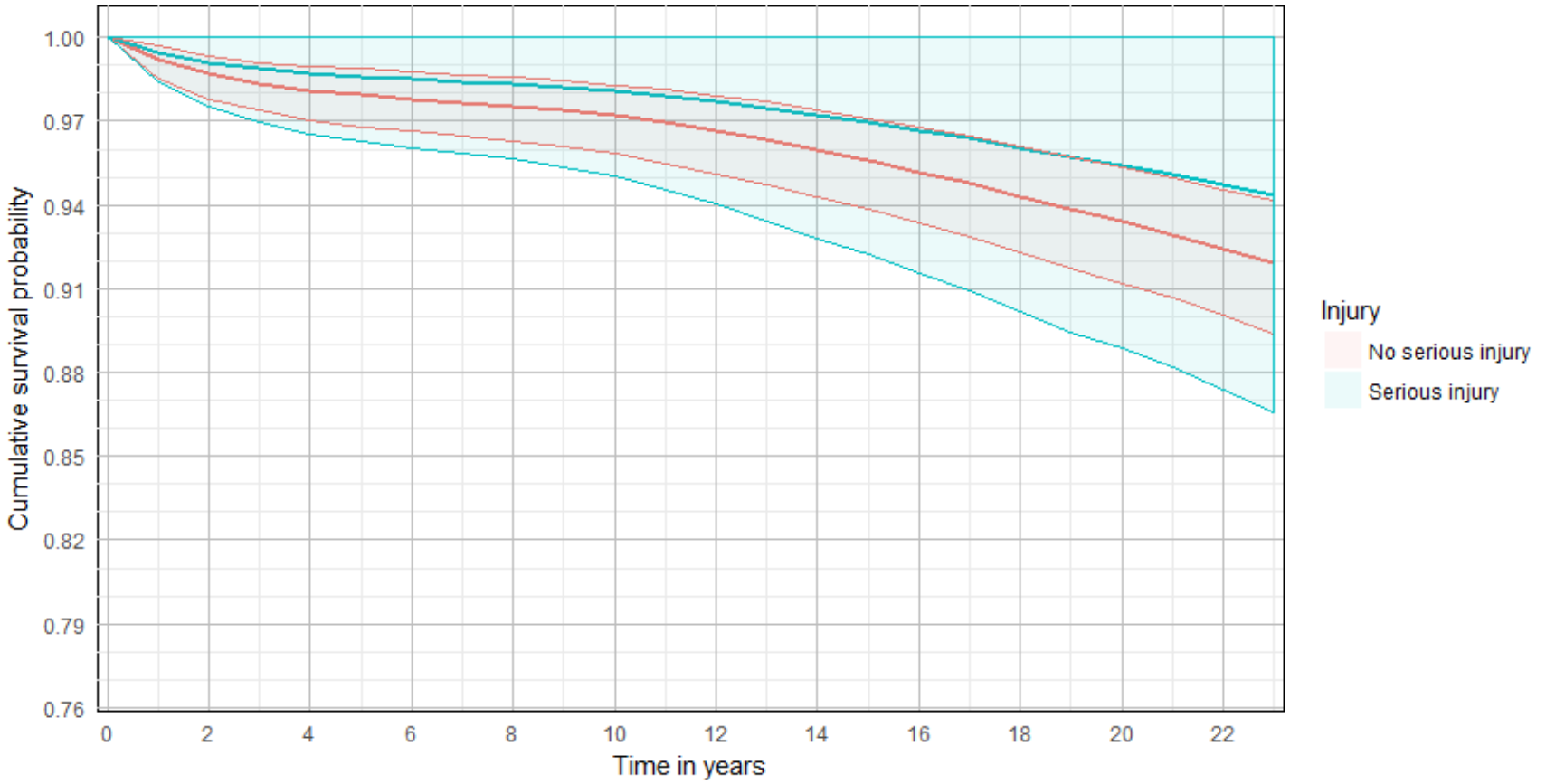

Figure 4.3 Cumulative survival probability from risk of diabetes among a cohort of the 1988 Spitak earthquake survivors during 1988-2012, by serious injury

[a] All the measures are standardized for age, age squared, gender, education in years, standard of living at the time of earthquake, place of residence (region), housing damage, and earthquake-related death of a family member among the total baseline population

[b] Point estimates and 95% confidence limits are based on the 50th, 2.5th, and 97.5th percentile obtained from 10,000 bootstrap sample estimates

Table 4.5 Marginal absolute effects of earthquake caused serious injury on cumulative survival probability from risk of diabetes among a cohort of the 1988 Spitak earthquake survivors

	Difference in cumulative survival probability [a] (95%CI) [b]	
	Earthquake-related injury	
Time (year)	**No serious injury**	**Serious injury**
1	Reference	0.002 (−0.008, 0.011)
2	Reference	0.004 (−0.012, 0.016)
3	Reference	0.005 (−0.014, 0.020)
5	Reference	0.007 (−0.017, 0.025)
10	Reference	0.009 (−0.023, 0.032)
15	Reference	0.013 (−0.034, 0.047)
23	Reference	0.024 (−0.056, 0.084)

[a] All the measures are standardized for age, age squared, gender, education in years, standard of living at the time of earthquake, place of residence (region), housing damage, and earthquake-related death of a family member among the total baseline population

[b] Point estimates and 95% confidence limits are based on the 50th, 2.5th, and 97.5th percentile obtained from 10,000 bootstrap sample estimates

Interventional approach for path-specific estimation of the effect of housing damage on cumulative probability of remaining diabetes-free

Details on the path-specific estimation of the effect of housing damage on cumulative probability of remaining diabetes-free in respect to receiving permanent housing and preventing job loss after the earthquake are provided in Table 4.6. A large proportion of the total effect of moderate housing damage (compared to no housing damage) on the cumulative probability of remaining diabetes-free. The mediating effect of receiving housing damage mediated through receiving a permanent housing was small, but in the opposite direction of the total effect of the housing damage on diabetes.

A similar pattern was observed for the effect of moderate housing damage mediated through job loss. The effect of moderate housing damage on risk of diabetes mediated by dependence of receiving permanent housing and loss of job were consistent with the total effect of housing damage on risk of diabetes. The direct effect of total destruction of housing (not through receiving permanent housing and job loss) on diabetes risk constituted a major component of the total effect of housing loss on diabetes risk. The effect mediated through receiving permanent housing was relatively small and in the opposite direction to the total effect of total housing destruction on diabetes. The effect mediated through job loss was also relatively small; nevertheless, it had the same direction as the total effect. Dependence between receiving permanent housing and loss of job mediated the effect of total housing destruction on risk of diabetes and it increased the risk of diabetes among those who experienced total destruction of housing.

The results of sensitivity analyses further adjusting for the reported smoking status and BMI of the participant at the time of the earthquake in the models for the mediators, as well as the outcome did affect the overall effect of housing damage on risk of diabetes. However, for both moderate housing damage and its total destruction, the effect mediated through receiving permanent housing, job loss, and the dependence between them were attenuated. Table A-4.4 presents the results of the sensitivity analyses applying interventional approach for estimating the path-specific effect of housing damage on the cumulative probability of remaining diabetes-free.

Table 4.6 Marginal absolute effects of earthquake-related housing damage on cumulative survival probability from risk of diabetes among a cohort of the 1988 Spitak earthquake survivors

		Difference in cumulative survival probability [a] (95%CI) [b]	
		Housing damage	
Time (year)	**Effect measure**	**Moderate damage**	**Total destruction**
1	IDE not through $\{M_1, M_2\}$	0.001 (−0.006, 0.008)	−0.004 (−0.015, 0.006)
1	IIE through M_1	0.000 (−0.006, 0.005)	0.000 (−0.007, 0.008)
1	IIE through M_2	−0.001 (−0.010, 0.008)	0.000 (−0.008, 0.007)
1	IIE through the dependence of M_2 on M_1	0.001 (−0.012, 0.015)	0.001 (−0.014, 0.014)
2	IDE not through $\{M_1, M_2\}$	0.002 (−0.008, 0.011)	−0.006 (−0.023, 0.008)
2	IIE through M_1	−0.001 (−0.008, 0.007)	0.001 (−0.008, 0.010)
2	IIE through M_2	−0.001 (−0.013, 0.009)	−0.001 (−0.009, 0.009)
2	IIE through the dependence of M_2 on M_1	0.002 (−0.014, 0.019)	0.001 (−0.018, 0.018)
3	IDE not through $\{M_1, M_2\}$	0.002 (−0.009, 0.013)	−0.007 (−0.027, 0.009)
3	IIE through M_1	−0.001 (−0.008, 0.007)	0.001 (−0.009, 0.011)
3	IIE through M_2	−0.002 (−0.015, 0.010)	−0.001 (−0.011, 0.009)
3	IIE through the dependence of M_2 on M_1	0.003 (−0.016, 0.021)	0.001 (−0.019, 0.020)
5	IDE not through $\{M_1, M_2\}$	0.002 (−0.012, 0.017)	−0.009 (−0.035, 0.012)
5	IIE through M_1	−0.001 (−0.011, 0.009)	0.001 (−0.012, 0.014)
5	IIE through M_2	−0.002 (−0.018, 0.012)	−0.001 (−0.014, 0.012)
5	IIE through the dependence of M_2 on M_1	0.004 (−0.019, 0.027)	0.001 (−0.023, 0.026)
10	IDE not through $\{M_1, M_2\}$	0.003 (−0.014, 0.021)	−0.012 (−0.045, 0.014)
10	IIE through M_1	−0.002 (−0.013, 0.010)	0.001 (−0.014, 0.017)
10	IIE through M_2	−0.004 (−0.022, 0.013)	−0.002 (−0.017, 0.014)
10	IIE through the dependence of M_2 on M_1	0.006 (−0.022, 0.034)	0.002 (−0.026, 0.032)
15	IDE not through $\{M_1, M_2\}$	0.005 (−0.020, 0.031)	−0.019 (−0.068, 0.020)
15	IIE through M_1	−0.002 (−0.017, 0.013)	0.002 (−0.017, 0.022)
15	IIE through M_2	−0.005 (−0.028, 0.016)	−0.003 (−0.022, 0.017)
15	IIE through the dependence of M_2 on M_1	0.009 (−0.025, 0.043)	0.004 (−0.033, 0.042)
23	IDE not through $\{M_1, M_2\}$	0.008 (−0.032, 0.050)	−0.033 (−0.111, 0.033)
23	IIE through M_1	−0.003 (−0.024, 0.018)	0.003 (−0.023, 0.032)
23	IIE through M_2	−0.009 (−0.039, 0.019)	−0.004 (−0.031, 0.022)
23	IIE through the dependence of M_2 on M_1	0.013 (−0.034, 0.061)	0.007 (−0.045, 0.060)

IDE: Interventional direct effect; IIE: Interventional indirect effect.

[a] All the measures are standardized for age, age squared, gender, education in years, standard of living at the time of earthquake, place of residence (region), earthquake-related death of a family member, and serious injury among the total baseline population

[b] Point estimates and 95% confidence limits are based on the 50th, 2.5th, and 97.5th percentile obtained from 10,000 bootstrap sample estimates

4.4 Discussion

This study examined the independent impacts of the death of a family member, sustaining a severe injury, and damage to one's housing on diabetes risk in a cohort of individuals with differential exposure levels to the 1988 Spitak earthquake. The study further decomposed the total effect of housing damage on risk of incident diabetes into specific pathways defined with respect to potential mediators: receiving permanent housing and job loss in the aftermath of the earthquake.

We found an increased risk of diabetes among those whose housing was totally destroyed. Nevertheless, the effect of moderate housing damage was not differentiable from the null effect. While no longitudinal study has evaluated the effect of housing damage due to earthquake on diabetes risk, studies have documented association between material loss, including housing damage, and PTSD (Armenian et al., 2000; Zhang et al., 2012), and PTSD has been found to be associated with higher diabetes incidence (Roberts et al., 2015). Additionally, studies have shown a positive link between elevated stress and diabetes risk(Kelly & Ismail, 2015). The results of the current study are consistent with those findings.

As with all-cause mortality, this study did not reveal any association between the death of a family member and diabetes risk. Considering that the death of a family member is a stressful event and that such events are directly and indirectly (through PTSD) associated with diabetes, our findings regarding the effect of the death of a family member and risk diabetes appear somewhat inconsistent (Kelly & Ismail, 2015; Roberts et al., 2015). Nevertheless, mechanisms and pathways other than elevated stress and PTSD might have qualitatively different effects, resulting in an overall null effect. Further studies assessing various pathways and potential mechanisms of action connecting the death of a family member and health outcomes will expand our understanding of the causal structure.

The study found a moderate negative association between severe injury and risk of diabetes. It is possible that at least some of the observed effect could be attributable to survival bias. We anticipated that severe injury was associated with certain individual characteristics at baseline which were, in turn, associated with diabetes risk. The attenuation of the effect of severe injury on risk of diabetes in the sensitivity analyses that further adjusted for smoking habit and BMI of participants at the baseline confirmed our anticipated results. Although adjustment for smoking habit and BMI decreased the observed effect, the results were still suggestive of a negative association between severe injury and risk of diabetes. Such an association could be due, in part, due to selection bias and residual confounding.

Path-specific decomposition of the effect of housing damage on diabetes risk with respect to receiving permanent housing and job loss after the earthquake showed that most of the effect of housing damage on diabetes was due to its direct effect. As expected, the effect of housing damage through job loss increased diabetes risk, while the effect mediated through receiving a permanent housing had the opposite effect. The effect mediated through dependence of the mediators, receiving permanent housing and job loss, was notable for total destruction of housing as well as for moderate housing damage.

To the best of our knowledge, this is the first study to investigate the independent effect of selected earthquake exposures on diabetes risk and quantify the path-specific contribution of housing damage on diabetes risk. The data obtained from a prospective study and modern causal inference approaches allowed rigorous examination of the effect of each exposure of interest on the outcome. The interventional approach for path-specific causal mediation analyses offered a platform to assess multiple mediators, to relax the no-exposure-induced mediator-outcome confounding assumption, and to incorporate exposure-mediator interaction.

The study had several methodological limitations. All measures were survey-based and self-reported and could be subject to information bias. Nevertheless, the prospective nature of the study and verification of the measurements in subsequent phases improved the quality of data and minimized the potential for information bias. Although we used modern causal techniques to estimate the effect of earthquake-related exposures on diabetes risk, the results could have been subject to potential biases including residual confounding. Potential selective survival bias introduced by the earthquake-related exposure and subsequent deaths could have biased the study results. Although such bias, if any, was likely to be small and toward the null, it could still have reduced the accuracy of our results.

Long-term follow-up assessments of disaster survivors could paint a better picture of the potential short- and long-term health ramifications of disasters and the factors connected to such outcomes. The findings of similar evaluations can contribute to early risk assessment, timely interventions, and planning for long-term initiatives. In addition, they can serve as resources upon which policymakers can draw while developing policies and strategies that target disaster areas and affected populations.

4.5 Appendix

Appendix figures

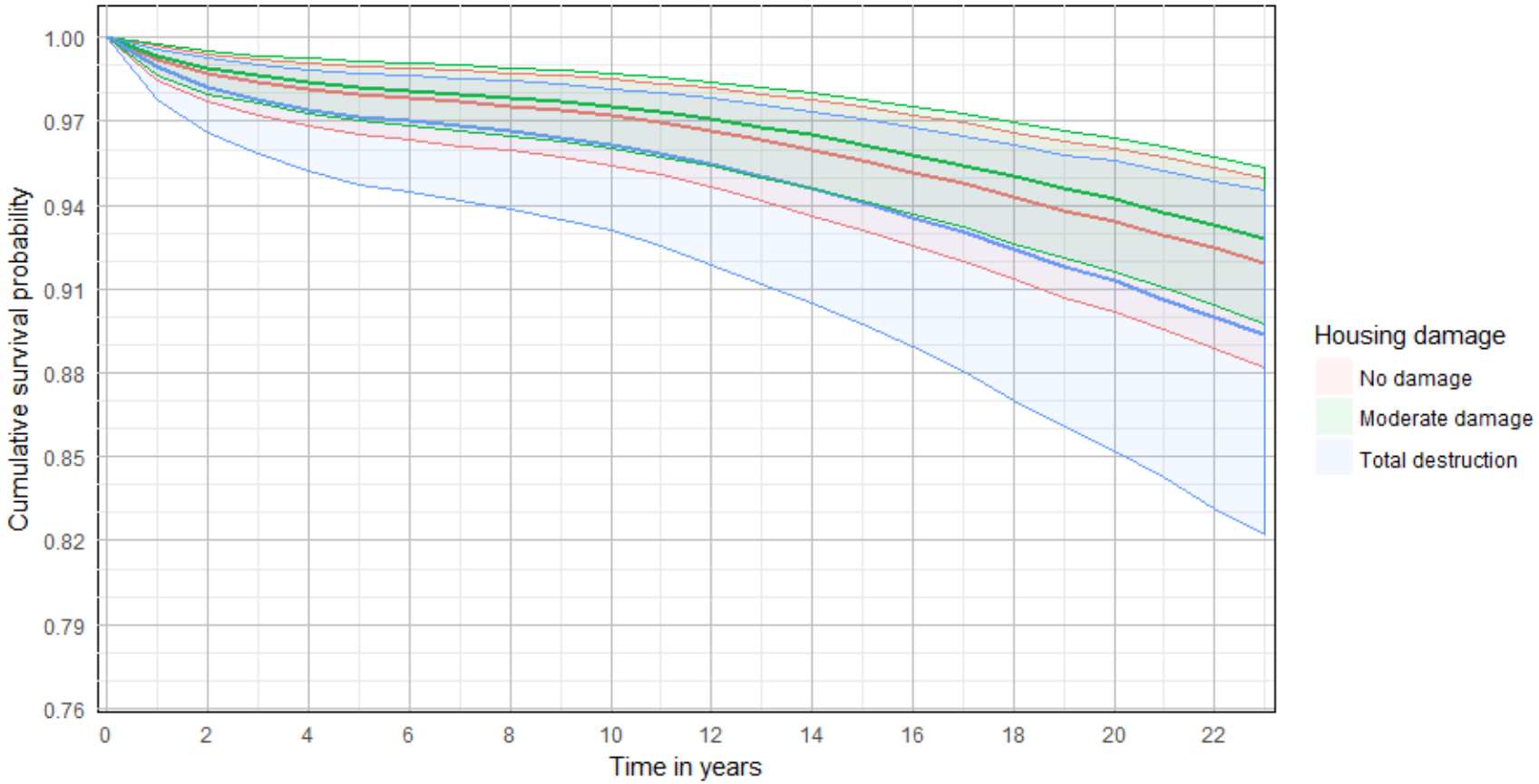

Figure A-4.1 Cumulative survival probability from risk of diabetes among a cohort of the 1988 Spitak earthquake survivors during 1988-2012, by housing damage levels

[a] All the measures are standardized for age, age squared, gender, education in years, smoking and drinking habits, standard of living at the time of earthquake, place of residence (region), earthquake-related death of a family member, and serious injury among the total baseline population

[b] Point estimates and 95% confidence limits are based on the 50th, 2.5th, and 97.5th percentile obtained from 10,000 bootstrap sample estimates

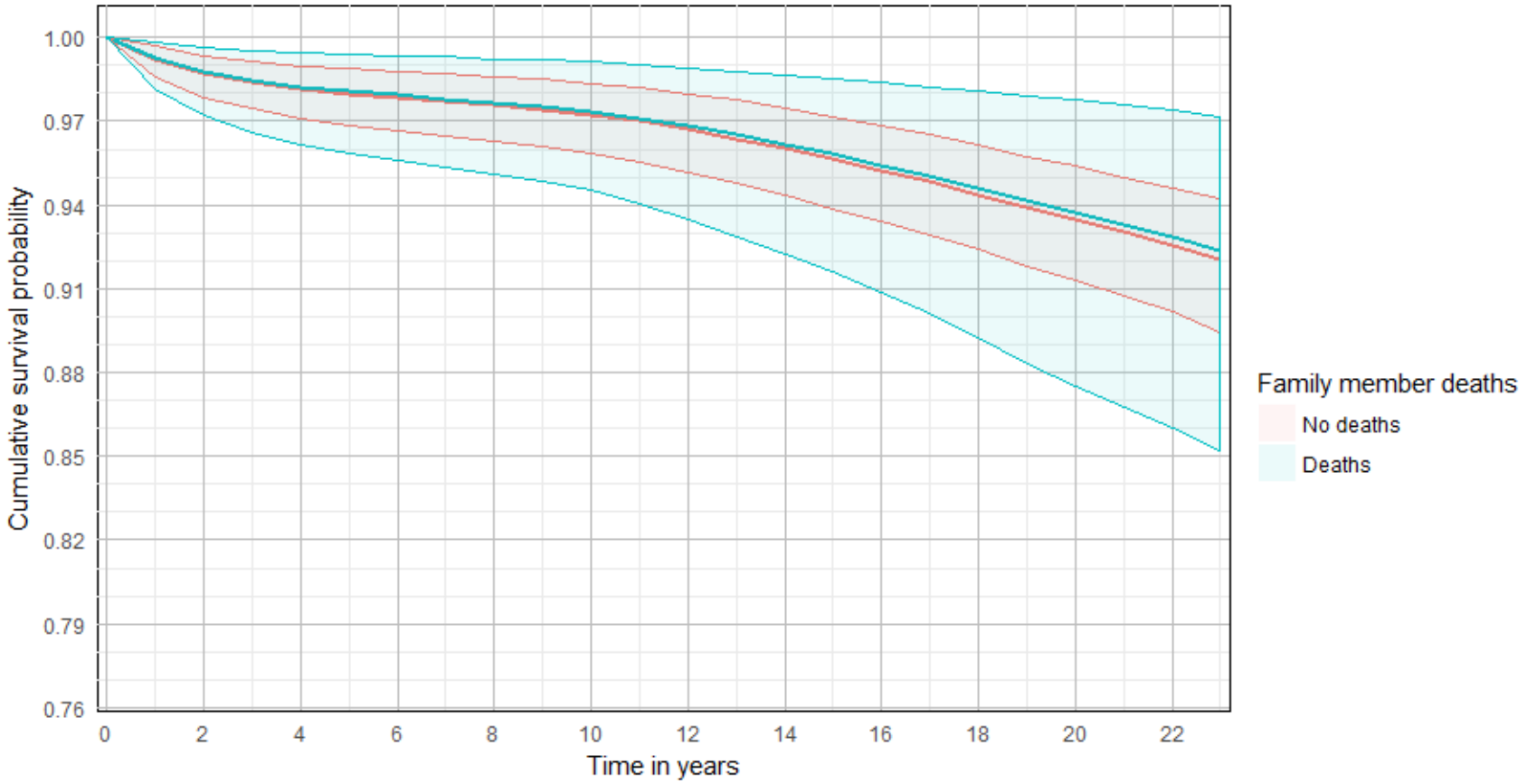

Figure A-4.2 Cumulative survival probability from risk of diabetes among a cohort of the 1988 Spitak earthquake survivors during 1988-2012, by family member death, using outcome data partially imputed by random survival forest algorithm

[a] All the measures are standardized for age, age squared, gender, education in years, smoking and drinking habits, standard of living at the time of earthquake, place of residence (region), housing damage, and serious injury among the total baseline population

[b] Point estimates and 95% confidence limits are based on the 50th, 2.5th, and 97.5th percentile obtained from 10,000 bootstrap sample estimates

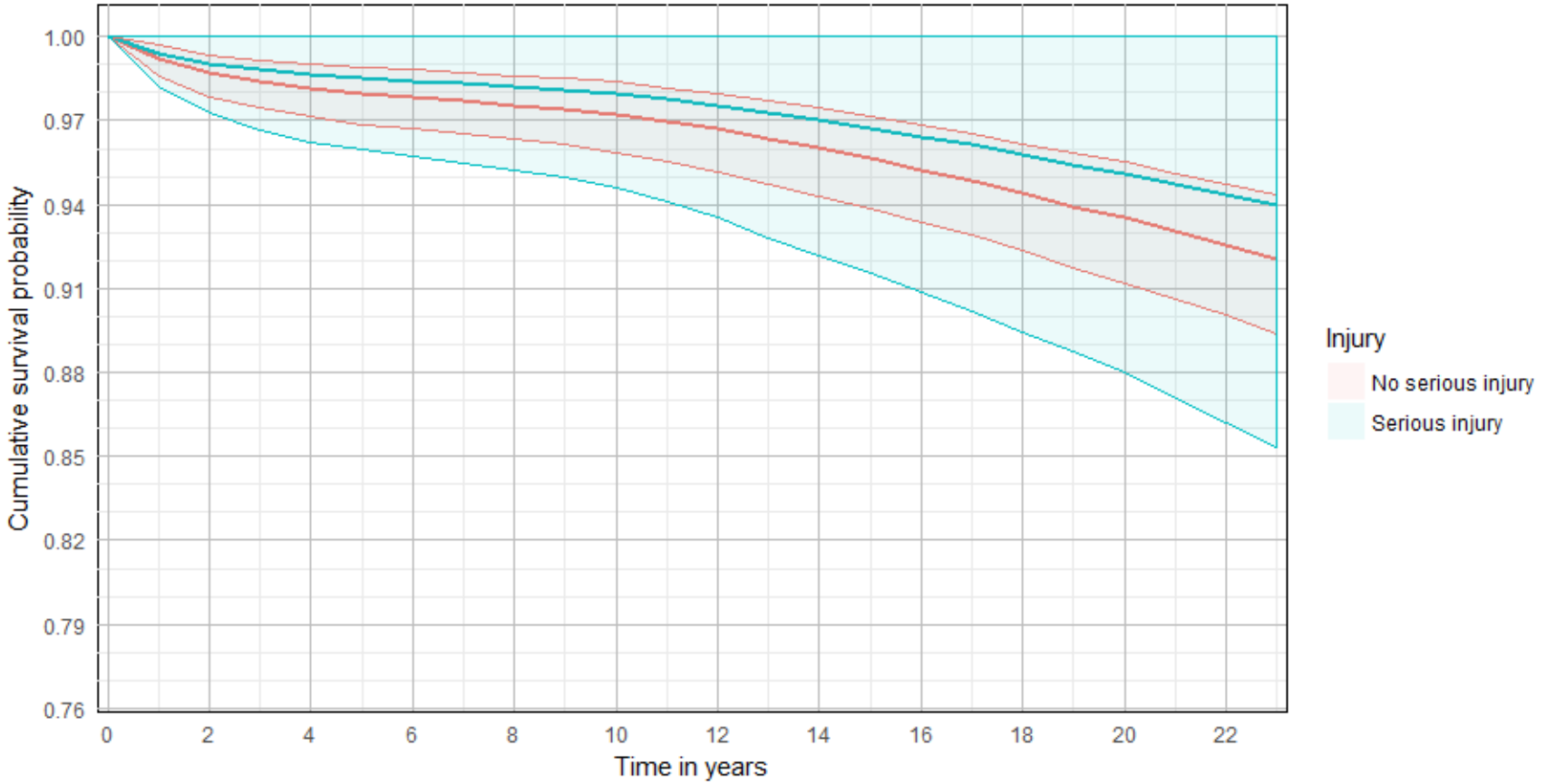

Figure A-4.3 Cumulative survival probability from risk of diabetes among a cohort of the 1988 Spitak earthquake survivors during 1988-2012, by injury status

[a] All the measures are standardized for age, age squared, gender, education in years, smoking and drinking habits, standard of living at the time of earthquake, place of residence (region), housing damage, and earthquake-related death of a family member among the total baseline population

[b] Point estimates and 95% confidence limits are based on the 50th, 2.5th, and 97.5th percentile obtained from 10,000 bootstrap sample estimates

Appendix tables

Table A-4.1 Marginal absolute effects of earthquake-related housing damage on cumulative survival probability from risk of diabetes among a cohort of the 1988 Spitak earthquake survivors

	Difference in cumulative survival probability [a] (95%CI) [b]		
	Housing damage		
Time (year)	**No damage**	**Moderate damage**	**Total destruction**
1	Reference	0.001 (−0.006, 0.009)	−0.003 (−0.015, 0.007)
2	Reference	0.002 (−0.008, 0.012)	−0.005 (−0.022, 0.009)
3	Reference	0.002 (−0.009, 0.014)	−0.006 (−0.026, 0.011)
5	Reference	0.003 (−0.012, 0.017)	−0.007 (−0.033, 0.013)
10	Reference	0.004 (−0.014, 0.023)	−0.010 (−0.043, 0.017)
15	Reference	0.006 (−0.020, 0.033)	−0.015 (−0.062, 0.025)
23	Reference	0.009 (−0.034, 0.054)	−0.026 (−0.104, 0.040)

[a] All the measures are standardized for age, age squared, gender, education in years, smoking and drinking habits, standard of living at the time of earthquake, place of residence (region), earthquake-related death of a family member, and serious injury among the total baseline population

[b] The point estimates and 95% confidence limits are based on the 50th, 2.5th, and 97.5th percentile obtained from 10,000 bootstrap sample estimates

Table A-4.2 Marginal absolute effects of earthquake-related death of a family member on cumulative survival probability from risk of diabetes among a cohort of the 1988 Spitak earthquake survivors

	Difference in cumulative survival probability [a] (95%CI) [b]	
	Earthquake-related death in the family member	
Time (year)	**No death**	**Death of a family member**
1	Reference	0.001 (−0.010, 0.008)
2	Reference	0.001 (−0.015, 0.012)
3	Reference	0.001 (−0.018, 0.014)
5	Reference	0.001 (−0.022, 0.018)
10	Reference	0.001 (−0.029, 0.024)
15	Reference	0.002 (−0.042, 0.034)
23	Reference	0.003 (−0.072, 0.058)

[a] All the measures are standardized for age, age squared, gender, education in years, smoking and drinking habits, standard of living at the time of earthquake, place of residence (region), housing damage, and serious injury among the total baseline population

[b] Point estimates and 95% confidence limits are based on the 50th, 2.5th, and 97.5th percentile obtained from 10,000 bootstrap sample estimates

Table A-4.3 Marginal absolute effects of earthquake caused serious injury on cumulative survival probability from risk of diabetes among a cohort of the 1988 Spitak earthquake survivors

	Difference in cumulative survival probability [a] (95%CI) [b]	
	Earthquake-related injury	
Time (year)	**No serious injury**	**Serious injury**
1	Reference	0.002 (−0.009, 0.011)
2	Reference	0.003 (−0.014, 0.015)
3	Reference	0.004 (−0.017, 0.019)
5	Reference	0.005 (−0.021, 0.024)
10	Reference	0.007 (−0.027, 0.032)
15	Reference	0.011 (−0.041, 0.047)
23	Reference	0.018 (−0.070, 0.083)

[a] All the measures are standardized for age, age squared, gender, education in years, smoking and drinking habits, standard of living at the time of earthquake, place of residence (region), housing damage, and earthquake-related death of a family member among the total baseline population

[b] Point estimates and 95% confidence limits are based on the 50th, 2.5th, and 97.5th percentile obtained from 10,000 bootstrap sample estimates

Table A-4.4 Marginal absolute effects of earthquake-related housing damage on cumulative survival probability from risk of diabetes among a cohort of the 1988 Spitak earthquake survivors

		Difference in cumulative survival probability [a] (95%CI) [b]	
		Housing damage	
Time (year)	**Effect measure**	**Moderate damage**	**Total destruction**
1	IDE not through {M_1 , M_2}	0.001 (−0.006, 0.008)	−0.004 (−0.015, 0.006)
1	IIE through M_1	0.000 (−0.005, 0.006)	0.000 (−0.007, 0.008)
1	IIE through M_2	0.000 (−0.005, 0.005)	0.000 (−0.007, 0.007)
1	IIE through the dependence of M_2 on M_1	0.000 (−0.011, 0.012)	0.000 (−0.014, 0.014)
2	IDE not through {M_1 , M_2}	0.001 (−0.008, 0.011)	−0.005 (−0.023, 0.008)
2	IIE through M_1	0.000 (−0.007, 0.007)	0.001 (−0.009, 0.009)
2	IIE through M_2	0.000 (−0.008, 0.007)	0.000 (−0.009, 0.009)
2	IIE through the dependence of M_2 on M_1	0.001 (−0.014, 0.015)	0.000 (−0.017, 0.018)
3	IDE not through {M_1 , M_2}	0.001 (−0.009, 0.013)	−0.007 (−0.028, 0.009)
3	IIE through M_1	0.001 (−0.008, 0.008)	0.001 (−0.009, 0.011)
3	IIE through M_2	−0.001 (−0.008, 0.007)	0.000 (−0.010, 0.010)
3	IIE through the dependence of M_2 on M_1	0.001 (−0.015, 0.017)	0.001 (−0.019, 0.020)
5	IDE not through {M_1 , M_2}	0.001 (−0.012, 0.016)	−0.009 (−0.035, 0.011)
5	IIE through M_1	0.001 (−0.010, 0.010)	0.001 (−0.012, 0.013)
5	IIE through M_2	0.000 (−0.010, 0.009)	0.000 (−0.012, 0.013)
5	IIE through the dependence of M_2 on M_1	0.001 (−0.018, 0.021)	0.001 (−0.023, 0.024)
10	IDE not through {M_1 , M_2}	0.002 (−0.016, 0.021)	−0.012 (−0.046, 0.015)
10	IIE through M_1	0.001 (−0.012, 0.013)	0.001 (−0.014, 0.016)
10	IIE through M_2	−0.001 (−0.012, 0.010)	0.001 (−0.014, 0.015)
10	IIE through the dependence of M_2 on M_1	0.002 (−0.021, 0.025)	0.001 (−0.027, 0.028)
15	IDE not through {M_1 , M_2}	0.003 (−0.022, 0.031)	−0.019 (−0.068, 0.021)
15	IIE through M_1	0.001 (−0.015, 0.016)	0.001 (−0.017, 0.020)
15	IIE through M_2	−0.001 (−0.015, 0.014)	0.001 (−0.018, 0.019)
15	IIE through the dependence of M_2 on M_1	0.003 (−0.027, 0.032)	0.002 (−0.032, 0.036)
23	IDE not through {M_1 , M_2}	0.006 (−0.036, 0.048)	−0.032 (−0.111, 0.034)
23	IIE through M_1	0.001 (−0.020, 0.023)	0.002 (−0.023, 0.029)
23	IIE through M_2	−0.002 (−0.023, 0.018)	0.001 (−0.025, 0.025)
23	IIE through the dependence of M_2 on M_1	0.005 (−0.038, 0.047)	0.004 (−0.042, 0.050)

IDE: Interventional direct effect; IIE: Interventional indirect effect.

[a] All the measures are standardized for age, age squared, gender, education in years, smoking and drinking habits, standard of living at the time of earthquake, place of residence (region), earthquake-related death of a family member, and serious injury among the total baseline population

[b] Point estimates and 95% confidence limits are based on the 50th, 2.5th, and 97.5th percentile obtained from 10,000 bootstrap sample estimates

Chapter 5. Comparison of hypothetical interventions to reduce post-earthquake morbidity and mortality

Abstract

Although findings from studies have been used to develop interventions targeting disaster survivors, no formal study has projected, simulated, and formally evaluated the impact of scientifically supported hypothetical interventions. The current study used the results and estimates from previous chapters to evaluate the impacts of hypothetical post-earthquake interventions on diabetes and mortality. The hypothetical interventions were based on providing permanent housing and preventing job loss in the aftermath of the 1988 Spitak earthquake. Each intervention was implemented under conservative to more optimistic scenarios. The results indicate that providing housing aid to those with total housing destruction was the most beneficial intervention to decrease the risk of all-cause mortality. Both providing housing aid and preventing job loss were effective in decreasing the risk of diabetes. The combination of both interventions found to be slightly more effective than each of the interventions alone. This study highlighted the importance of post-earthquake interventions on all-cause mortality and diabetes. It demonstrated the positive impact of providing housing aid and preventing job loss among earthquake survivors on the risk of all-cause mortality and diabetes. The findings of this study can be used to prioritize interventions in the aftermath of disasters. The results can also serve as a guide for public health practitioners working on disaster relief programs.

5.1 Introduction

Most research studies and literature to date have focused on the perspectives of the scientific community. Although earthquakes are unpredictable, preparedness can help mitigate and prevent their negative impact. High levels of preparedness require political will, technological

capabilities, long-term planning, and substantial financial resources. Limited resources, and lack of political will, especially in low- and middle-income countries, make populations and communities vulnerable.

Although preparedness may mitigate the harm caused by the disaster, it often does not entirely eliminate such outcomes. Moreover, from time to time an unusually severe disaster can surpass a country's preparedness and resilience, however extensive, causing significant damages. The Great East Japan earthquake of 2011 is an example of such an event. In the presence of damage from a disaster, post-disaster management and interventions targeting the affected communities become the primary sources for damage control aiming to mitigate adverse consequences.

Although findings from studies have been used to develop interventions targeting disaster survivors, no formal study has projected, simulated, and formally evaluated the impact of scientifically supported hypothetical interventions. This study utilized findings from the previous chapters to evaluate the impacts of hypothetical interventions on diabetes and mortality. Such a simulation study can provide useful information for decision making purposes; moreover, it can serve as an example encouraging future studies to consider similar policy relevant analysis.

5.2 Methods

Building on the results from Chapters 3 and 4, the current study evaluated the impact of two post-disaster interventions, providing permanent housing and preventing job loss, on the risk of diabetes and on mortality in a cohort of earthquake survivors. Each intervention was implemented under various scenarios, ranging from conservative to more optimistic scenarios. The interventions were designed by optimizing the distribution of the existing level of the desired factor (i.e. providing permanent housing or preventing job loss) or changing the distribution of

the positive post-disaster factor to a more desirable level using simulations. Description of the variables and their measurements are provided in Chapters 2, 3, and 4.

To assess the effect of providing permanent housing in the aftermath of the earthquake on mortality, we considered survival rates under the existing (observed) rates and distribution of provision of permanent housing as the reference situation. The hypothetical interventions included more effective allocation of the observed level of permanent housing aid in the aftermath of the earthquake, providing housing to only those whose housing were destroyed, providing housing to all with moderate or extensive housing damage, including those whose houses were totally destroyed, and finally, providing housing to all participants, regardless of the level of damage to their housing.

The hypothetical interventions targeting diabetes rates were based on providing permanent housing, or preventing job loss in the aftermath of the event. The hypothetical interventions with regards to the provision of housing were similar the ones defined for all-cause mortality. The interventions for job loss ranged from elimination of job loss among those with total destruction of housing, elimination of job loss among those with moderate housing damage or total destruction of housing, and the most optimistic scenario, prevention of all the instances of job loss during the post-earthquake period.

Statistical analysis

Using parameters obtained from the survival models applied on the bootstrap samples in Chapters 3 and 4, we implemented the g-computation algorithm by conducting Monte Carlo simulations to compute the potential outcomes under various hypothetical interventions. After obtaining the potential outcomes, we generated the risk of all-cause mortality and risk of

diabetes and constructed relative effect measures (i.e. risk ratio) in order to compare and evaluate the effectiveness of the hypothetical interventions relative to the observed situation.

5.3 Results

All-cause mortality

Among the 23,639 adults with differential levels of exposure to the 1988 Spitak earthquake, 19.1% experienced total destruction of their housing due to the earthquake, 44.2% sustained moderate damage to their housing, and 36.7% had minor to no damage to their housing. Among the entire sample, a total of 14.9% received new permanent housing in the aftermath of the disaster. The likelihood of receiving permanent housing was directly associated with the level of damage to the participants housing at the time of the earthquake. Among those whose housing was totally destroyed, 20.9% received permanent housing. The probabilities of receiving permanent housing among those with moderate and those with minor or no damage to their housing were 15.3% and 11.4%, respectively.

The 23-year risk of all-cause mortality among the participants under the observed level of housing aid (natural course) was 0.257 (95% CI=0.250, 0.264). The 23-year risk of all-cause mortality under optimized allocation of housing and allocation of housing to all those with total destruction of housing was 0.254 (95% CI=0.247, 0.260) and 0.254 (95% CI=0.248, 0.262) respectively. Providing housing to all participants with moderate housing damage and total destruction of housing yielded a 23-year risk of all-cause mortality of 0.248 (95% CI=0.240, 0.258) while the 23-year risk of all-cause mortality under the hypothetical intervention of providing housing to all of the study participants was 0.245 (95% CI=0.234, 0.257).

The return was the highest for providing housing to those with total housing destruction followed by those with moderate housing damage and finally, to those with minor to no housing damage. Nevertheless, the intervention (housing aid) remained moderately effective in reducing all-cause mortality among all aid recipients, regardless of their level of housing damage.

Table 5.1 Simulated 23-year cumulative risk of all-cause mortality under hypothetical interventions providing various levels of housing aid in a cohort exposed to the 1988 Spitak earthquake in Armenia (N=23,639)

Interventions	**Average percent who received housing (%)**	**23-year cumulative risk of all-cause mortality (95%CI)** [a]	**Risk ratio of mortality during a 23-year period (95%CI)** [a]
No housing aid at all	0.0%	0.259 (0.254, 0.266)	1.02 (1.00, 1.05)
Observed housing aid (natural course)	14.9%	0.257 (0.250, 0.264)	Referent
Redistribution of the observed level of housing aid (optimization)	14.9%	0.254 (0.247, 0.260)	0.99 (0.97, 1.02)
Housing aid to those with total destruction of housing	19.1%	0.254 (0.248, 0.262)	0.99 (0.97, 1.02)
Housing aid to those with moderate housing damage to total destruction of housing	63.3%	0.248 (0.240, 0.258)	0.96 (0.94, 1.01)
Housing aid to all	100%	0.245 (0.234, 0.257)	0.95 (0.92, 1.01)

[a] Point estimates and 95% confidence limits are based on the 50th, 2.5th, and 97.5th percentile obtained from 10,000 bootstrap sample estimates

Diabetes

Of the 23,137 participants in the sample, 19.0% had their housing totally destroyed, 44.2% had moderate damage to their housing, and 36.8% had minor to no damage to their housing. Fifteen percent of the sample received a permanent housing in the aftermath of the earthquake (21.1% of those with total housing destruction, 15.3% of those with moderate housing damage, and 11.5% with minor or no housing damage). Slightly more than a fifth of the sample, 22.3%, lost their jobs in 1988–1990 period and remained unemployed for more than 6 months despite looking for a job (29.8% of those with total housing destruction, 21.1% of those with moderate housing damage, and 19.9% with minor or no housing damage).

Under the observed level of permanent housing aid and job loss, the cumulative incidence of diabetes over the 23-year of the study follow-up was 0.079 (95% CI=0.062, 0.099). Under the hypothetical intervention of providing housing to all those who experienced total destruction of housing or moderate housing damage with no changes in the natural course of job loss, the cumulative risk of diabetes was 0.065 (95% CI=0.038, 0.100). Preventing job loss among those with total destruction of housing without changing the natural course of housing aid provided to the survivors yielded a 23-year diabetes risk of 0.080 (95% CI=0.061, 0.101). The risk of diabetes under the hypothetical intervention of providing housing aid and preventing job loss among those with moderate housing damage to total destruction of housing was 0.065 (95% CI=0.037, 0.103). Table 5.1 provides risk of diabetes under various hypothetical interventions targeting housing aid and job loss and offers comparisons of the estimated risk relative to the actual, observed risk of diabetes.

Table 5.2 Simulated 23-year cumulative risk of diabetes under hypothetical interventions targeting housing aid and job loss in a cohort exposed to the 1988 Spitak earthquake in Armenia (N=23,137)

Interventions	**Average percent who received housing (%)**	**Average percent who had job loss (%)**	**23-year cumulative risk of diabetes (95%CI)** [a]	**Risk ratio of diabetes during a 23-year period (95%CI)** [a]
No housing aid at all	0%	22.3%	0.082 (0.062, 0.105)	1.04 (0.93, 1.13)
Observed housing aid (natural course)	15.0%	22.3%	0.079 (0.062, 0.099)	Referent
Redistribution of the observed level of housing aid (optimization)	15.0%	22.3%	0.077 (0.057, 0.102)	0.98 (0.83, 1.14)
Housing aid to those with total destruction of housing	19.0%	22.3%	0.076 (0.054, 0.103)	0.96 (0.78, 1.17)
Housing aid to those with moderate housing damage to total destruction of housing	63.2%	22.3%	0.065 (0.038, 0.100)	0.82 (0.51, 1.23)
Housing aid to all	100%	22.3%	0.058 (0.022, 0.105)	0.74 (0.28, 1.31)
Preventing job loss among those with total destruction of housing	15.0%	16.7%	0.080 (0.061, 0.101)	1.01 (0.93, 1.09)
Preventing job loss among those with moderate housing damage to total destruction of housing	15.0%	7.3%	0.074 (0.056, 0.095)	0.93 (0.83, 1.04)
Preventing job loss among all	15.0%	0.0%	0.067 (0.050, 0.088)	0.85 (0.72, 0.98)
Providing housing aid and preventing job loss among those with total destruction of housing	19.0%	16.7%	0.077 (0.054, 0.107)	0.97 (0.78, 1.24)
Providing housing aid and preventing job loss among those with moderate housing damage to total destruction of housing	63.2%	7.3%	0.065 (0.037, 0.103)	0.82 (0.50, 1.25)
Providing housing aid and preventing job loss among all	100%	0.0%	0.056 (0.018, 0.107)	0.71 (0.23, 1.34)

[a] Point estimates and 95% confidence limits are based on the 50th, 2.5th, and 97.5th percentile obtained from 10,000 bootstrap sample estimates

5.4 Discussion

In this chapter we evaluated impact of hypothetical interventions targeting post-disaster circumstances, namely housing aid and job loss on the risk of all-cause mortality and diabetes in a cohort of survivors with various levels of exposure to the 1988 Spitak earthquake in Armenia.

In the previous chapters we assessed the impact of various earthquake-related factors, including their controlled direct effect not through some of the post-disaster factors as well as the indirect effect of those factors mediated through post-earthquake events such as job loss or housing aid. While those findings expand our understanding of the impact of earthquake exposure and the role of post-disaster events on all-cause mortality and diabetes, the findings do not directly translate policy changes and actions. The evaluation of the hypothetical interventions was an attempt to fill the gap between scientific research and its practical implications and make the research findings more accessible to policy makers and public health practitioners.

The results indicate that providing housing aid to those with total housing destruction was the most beneficial intervention to decrease the risk of all-cause mortality. Although providing housing was effective in decreasing all-cause mortality, aid to those who experienced total destruction of their housing showed the highest gain relative to the proportion who received the intervention. These findings highlight the effectiveness of post-disaster interventions on all-cause mortality and emphasizes the importance of strategic allocation of limited resources.

The interventions of housing aid and job loss prevention appeared to have some positive impact on reducing diabetes risk. The results indicated that providing housing to those whose housing

was totally destroyed was slightly more effective in decreasing the risk of diabetes than the intervention of preventing job loss among the same population. Combining both interventions did not offer further reduction in diabetes risk compared to providing housing in the aftermath of the event alone. Nevertheless, preventing job loss could still remain an effective intervention, as it might be more practical and require fewer resources than providing housing to disaster survivors. Moreover, preventing job loss might offer health and economic benefits.

The study had several limitations. The hypothetical interventions and their impacts were simulated using models which could be a simplification and abstraction of the complex causal relationships observed in the real world. Hence, they might not have captured some important mechanisms that disaster and post-disaster events might lead to diabetes or mortality. The parameters supplied to the models were obtained from studies presented in Chapters 3 and 4, and were subject to most of the limitations described in those chapters, including confounding and measurement error. Additionally, the evaluation of interventions targeting risk of diabetes were conducted in a closed cohort where participants were assumed not to be subject to loss to competing risks. Further research should assess the effect of additional interventions and their combinations, and also assess their impact on individuals with various sociodemographic characteristics.

This study highlighted the importance of post-earthquake events on two facets of health outcomes, all-cause mortality and diabetes. It demonstrated the positive impact of providing housing aid and preventing job loss among earthquake survivors on the risk of all-cause mortality and diabetes. The study also illustrated the usefulness of computer models and simulations for evaluation of various hypothetical interventions and effective allocation of existing resources in the aftermath of disasters.

References

Adler, A. (1943). Neuropsychiatric complications in victims of Boston's Cocoanut Grove disaster. *Journal of the American Medical Association*, *123*(17), 1098.

American Psychiatric Association. (1980). Diagnostic and statistical manual of mental disorders (3rd Edition). In *Diagnostic and Statistical Manual of Mental Disorders 3rd Edition*.

American Psychiatric Association. (1987). *Diagnostic and Statistical Manual of Mental Disorders, Third Edition, Revised (DSM-III-R)*. Washington DC: American Psychiatric Association.

Andreasen, N. C. (2010). Posttraumatic stress disorder: A history and a critique. *Annals of the New York Academy of Sciences*, *1208*(1), 67–71.

Aoki, T., Fukumoto, Y., Yasuda, S., Sakata, Y., Ito, K., Takahashi, J., … Shimokawa, H. (2012). The Great East Japan Earthquake Disaster and cardiovascular diseases. *European Heart Journal*, *33*(22), 2796–2803.

Armenian, H. K., Melkonian, A. K., & Hovanesian, A. P. (1998). Long term mortality and morbidity related to degree of damage following the 1988 earthquake in Armenia. *American Journal of Epidemiology*, *148*(11), 1077–1084.

Armenian, H. K., Melkonian, A., Noji, E. K., & Hovanesian, a P. (1997). Deaths and injuries due to the earthquake in Armenia: a cohort approach. *International Journal of Epidemiology*, *26*(4), 806–813.

Armenian, H. K., Morikawa, M., Melkonian, a K., Hovanesian, a P., Haroutunian, N., Saigh, P. A., … Akiskal, H. S. (2000). Loss as a determinant of PTSD in a cohort of adult survivors of the 1988 earthquake in Armenia: implications for policy. *Acta Psychiatrica Scandinavica*, *102*(1), 58–64.

Armenian, H. K., Morikawa, M., Melkonian, A. K., Hovanesian, A., Akiskal, K., & Akiskal, H. S. (2002). Risk factors for depression in the survivors of the 1988 earthquake in Armenia. *J*

Urban Health, *79*(3), 373–382.

Armenian, H. K., Noji, E. K., & Oganesian, A. P. (1992). A case-control study of injuries arising from the earthquake in Armenia, 1988. *Bulletin of the World Health Organization*, *70*(2), 251–257.

Baron, R. M., & Kenny, D. a. (1986). The moderator-mediator variable distinction in social psychological research: conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology*, *51*(6), 1173–1182.

Beristianos, M. H., Yaffe, K., Cohen, B., & Byers, A. L. (2016). PTSD and Risk of Incident Cardiovascular Disease in Aging Veterans. *The American Journal of Geriatric Psychiatry : Official Journal of the American Association for Geriatric Psychiatry*, *24*(3), 192–200.

Binder, S., & Sanderson, L. M. (1987). The role of the epidemiologist in natural disasters. *Annals of Emergency Medicine*, *16*(9), 1081–1084.

Boscarino, J. A. (2006). Posttraumatic stress disorder and mortality among U.S. Army veterans 30 years after military service. *Annals of Epidemiology*, *16*(4), 248–256.

Boyko, E. J., Jacobson, I. G., Smith, B., Ryan, M. A. K., Hooper, T. I., Amoroso, P. J., … Millennium Cohort Study Team. (2010). Risk of diabetes in U.S. military service members in relation to combat deployment and mental health. *Diabetes Care*, *33*(8), 1771–1777.

Brady, K. T., Killeen, T. K., Brewerton, T., & Lucerini, S. (2000). Comorbidity of psychiatric disorders and posttraumatic stress disorder. *The Journal of Clinical Psychiatry*, *61 Suppl 7*(suppl 7), 22–32.

Breiman, L. (2001). Random forests. *Machine Learning*.

Cardena, E., & Spiegel, D. (1993). Dissociative reactions to the San Francisco Bay Area earthquake of 1989. *American Journal of Psychiatry*, *150*(3), 474–478.

Chen, C. H., Tan, H. K. L., Liao, L. R., Chen, H. H., Chan, C. C., Cheng, J. J. S., … Lu, M. L. (2007). Long-term psychological outcome of 1999 Taiwan earthquake survivors: a survey

of a high-risk sample with property damage. *Comprehensive Psychiatry*, *48*(3), 269–275.

Clauw, D. J., Engel, C. C., Aronowitz, R., Jones, E., Kipen, H. M., Kroenke, K., … Wessely, S. (2003). Unexplained symptoms after terrorism and war: an expert consensus statement. *Journal of Occupational and Environmental Medicine / American College of Occupational and Environmental Medicine*, *45*(10), 1040–1048.

de la Fuente, R. (1990). The Mental Health Consequences of the 1985 Earthquakes in Mexico. *International Journal of Mental Health*, *19*(2), 21–29.

Dell'OSso, L., Carmassi, C., Massimetti, G., Conversano, C., Daneluzzo, E., Riccardi, I., … Rossi, A. (2011). Impact of traumatic loss on post-traumatic spectrum symptoms in high school students after the L'Aquila 2009 earthquake in Italy. *Journal of Affective Disorders*, *134*(1–3), 59–64.

Dell'Osso, L., Carmassi, C., Massimetti, G., Daneluzzo, E., Di Tommaso, S., & Rossi, a. (2011). Full and partial PTSD among young adult survivors 10 months after the L'Aquila 2009 earthquake: gender differences. *Journal of Affective Disorders*, *131*(1–3), 79–83.

Demirchyan, A., Goenjian, A. K., & Khachadourian, V. (2015). Factor Structure and Psychometric Properties of the Posttraumatic Stress Disorder (PTSD) Checklist and DSM-5 PTSD Symptom Set in a Long-Term Postearthquake Cohort in Armenia. *Assessment*, *22*(5), 594–606.

Dirkzwager, A. J. E., van der Velden, P. G., Grievink, L., & Yzermans, C. J. (2007). Disaster-related posttraumatic stress disorder and physical health. *Psychosomatic Medicine*, *69*(5), 435–440.

Dobson, A. J., Alexander, H. M., Malcolm, J. A., Steele, P. L., & Miles, T. A. (1991). Heart attacks and the Newcastle earthquake. *Medical Journal of Australia*.

Efron, B. (1979). Bootstrap Methods: Another Look at the Jackknife. *The Annals of Statistics*.

Fisher, R. A. (1935). The design of experiments. In *The design of experiments*.

Foa, E. B., Stein, D. J., & McFarlane, A. C. (2006). Symptomatology and psychopathology of mental health problems after disaster. *Journal of Clinical Psychiatry*, *67*(SUPPL. 2), 15–25.

Goenjian, A K, Pynoos, R. S., Steinberg, A. M., Najarian, L. M., Asarnow, J. R., Karayan, I., … Fairbanks, L. A. (1995). Psychiatric comorbidity in children after the 1988 earthquake in Armenia. *J Am Acad Child Adolesc Psychiatry*, *34*(9), 1174–1184.

Goenjian, A K, Steinberg, A. M., Najarian, L. M., Fairbanks, L. A., Tashjian, M., & Pynoos, R. S. (2000). Prospective study of posttraumatic stress, anxiety, and depressive reactions after earthquake and political violence. *The American Journal of Psychiatry*, *157*(6), 911–916.

Goenjian, Armen K., Khachadourian, V., Armenian, H., Demirchyan, A., & Steinberg, A. M. (2018). Posttraumatic Stress Disorder 23 Years After the 1988 Spitak Earthquake in Armenia. *Journal of Traumatic Stress*, *31*(1), 47–56.

Goldmann, E., & Galea, S. (2014). Mental health consequences of disasters. *Annual Review of Public Health*, *35*(October), 169–183.

Gradus, J. L., Farkas, D. K., Svensson, E., Ehrenstein, V., Lash, T. L., Milstein, A., … Sørensen, H. T. (2015). Associations between stress disorders and cardiovascular disease events in the Danish population. *BMJ Open*, *5*(12), e009334.

Green, B. L., Lindy, J. D., Grace, M. C., Gleser, G. C., Leonard, A. C., Korol, M., & Winget, C. (1990). Buffalo Creek survivors in the second decade: stability of stress symptoms. *American Journal of Orthopsychiatry*, *60*(1), 43–54.

Greenland, S., Robins, J. M., & Pearl, J. (1999). Confounding and Collapsibility in Causal Inference. *Statistical Science*, *14*(1), 29–46.

Grinker, R. R. (1944). TREATMENT OF WAR NEUROSES. *Journal of the American Medical Association*, *126*(3), 142.

Guo, J., He, H., Qu, Z., Wang, X., & Liu, C. (2017). Post-traumatic stress disorder and depression among adult survivors 8 years after the 2008 Wenchuan earthquake in China.

Journal of Affective Disorders, *210*, 27–34.

Hadjian, A. H. (1993). The Spitak, Armenia earthquake of 7 December 1988 — why so much destruction. *Soil Dynamics and Earthquake Engineering*, *12*(1), 1–24.

Hafeman, D. M., & Schwartz, S. (2009). Opening the black box: A motivation for the assessment of mediation. *International Journal of Epidemiology*, *38*(3), 838–845.

Halpern, J. Y. (2000). Axiomatizing Causal Reasoning. *Journal of Artificial Intelligence Research*.

Hastie, T. J., & Tibshirani, R. J. (1990). Generalized Additive Models. *Monographs on Statistics and Applied Probability*. New York: Chapman and Hall.

Hata, M., Akiyama, K., Wakui, S., Takasaka, A., & Shiono, M. (2012). Rapid aortic enlargement requiring surgery in patients with chronic type B aortic dissection. *The Thoracic and Cardiovascular Surgeon*, *60 Suppl 2*, e31-2.

Hernán, M. A., Brumback, B., & Robins, J. M. (2000). Marginal structural models to estimate the causal effect of zidovudine on the survival of HIV-positive men. *Epidemiology (Cambridge, Mass.)*, *11*(5), 561–570.

Hernán, M. A., Brumback, B., & Robins, J. M. (2001). Marginal Structural Models to Estimate the Joint Causal Effect of Nonrandomized Treatments. *Journal of the American Statistical Association*, *96*(454), 440–448.

Hernán, M. A., & Robins, J. M. (2006). Estimating causal effects from epidemiological data. *Journal of Epidemiology and Community Health*, *60*(7), 578–586.

Holen, A. (1991). A longitudinal study of the occurrence and persistence of post-traumatic health problems in disaster survivors. *Stress Medicine*, *7*(1), 11–17.

Holland, P. W. (1988). Causal Inference, Path Analysis, and Recursive Structural Equations Models. *Sociological Methodology*, *18*(1), 449.

Hsu, C.-H., & Taylor, J. M. G. (2009). Nonparametric comparison of two survival functions with

dependent censoring via nonparametric multiple imputation. *Statistics in Medicine*, *28*(3), 462–475.

Hull, A. M., Alexander, D. A., & Klein, S. (2002). Survivors of the Piper Alpha oil platform disaster: long-term follow-up study. *The British Journal of Psychiatry : The Journal of Mental Science*, *181*(5), 433–438.

Hyman, H. (1955). *Survey design and analysis: Principles, cases, and procedures*. New York: Free Press.

Ishwaran, H., & Kogalur, U. B. (2007). Random Survival Forests for R. *R News*, *7*(2), 25–31.

Ishwaran, H., Kogalur, U. B., Blackstone, E. H., & Lauer, M. S. (2008). Random survival forests. *Annals of Applied Statistics*, *2*(3), 841–860.

Jacobson, I. G., Ryan, M. A. K., Hooper, T. I., Smith, T. C., Amoroso, P. J., Boyko, E. J., … Bell, N. S. (2008). Alcohol use and alcohol-related problems before and after military combat deployment. *JAMA*, *300*(6), 663–675.

Jordan, H. T., Miller-Archie, S. A., Cone, J. E., Morabia, A., & Stellman, S. D. (2011). Heart disease among adults exposed to the September 11, 2001 World Trade Center disaster: results from the World Trade Center Health Registry. *Preventive Medicine*, *53*(6), 370–376.

Kamoi, K., Tanaka, M., Ikarashi, T., & Miyakoshi, M. (2006). Effect of the 2004 Mid-Niigata Prefecture earthquake on home blood pressure measurement in the morning in type 2 diabetic patients. *Clinical and Experimental Hypertension (New York, N.Y. : 1993)*, *28*(8), 719–729.

Kantarci, G., Vanholder, R., Tuglular, S., Akin, H., Ko??, M., ??zener, ??etin, & Akoglu, E. (2002). Acute renal failure due to crush syndrome during Marmara earthquake. *American Journal of Kidney Diseases*, *40*(4), 682–689.

Kario, K., Matsuo, T., Shimada, K., & Pickering, T. G. (2001). Factors associated with the occurrence and magnitude of earthquake-induced increases in blood pressure. *The*

American Journal of Medicine, *111*(5), 379–384.

Kario, K., & Ohashi, T. (1997). Increased coronary heart disease mortality after the Hanshin-Awaji earthquake among the older community on Awaji Island. Tsuna Medical Association. *Journal of the American Geriatrics Society*, *45*(5), 610–613.

Kazama, M., & Noda, T. (2012). Damage statistics (Summary of the 2011 off the Pacific Coast of Tohoku Earthquake damage). *Soils and Foundations*, *52*(5), 780–792.

Kellet, J. (2014). *Disaster Risk Reduction Makes Development Sustainable* (D. Hillier, Ed.). United Nations Development Programme.

Kelly, S. J., & Ismail, M. (2015). Stress and type 2 diabetes: a review of how stress contributes to the development of type 2 diabetes. *Annual Review of Public Health*, *36*, 441–462.

Kenny, D. A. (2008). Reflections on Mediation. *Organizational Research Methods*, *11*(2), 353–358.

Khachadourian, V. (2011). Risk Factors and Outcomes of Post-Traumatic Stress Disorder and Comorbid Mental Disorders among 1988 Spitak Earthquake Survivors. *College of Health Sciences, American University of Armenia, Yerevan*.

Khachadourian, V., Armenian, H., Demirchyan, A., Melkonian, A., & Hovanesian, A. (2016). A post-earthquake psychopathological investigation in Armenia: methodology, summary of findings, and follow-up. *Disasters*, *40*(3), 518–533.

Khachadourian, V., Armenian, H. K., Demirchyan, A., & Goenjian, A. (2015). Loss and psychosocial factors as determinants of quality of life in a cohort of earthquake survivors. *Health and Quality of Life Outcomes*, *13*(1), 13.

Koenen, K. C., Sumner, J. A., Gilsanz, P., Glymour, M. M., Ratanatharathorn, A., Rimm, E. B., … Kubzansky, L. D. (2017). Post-traumatic stress disorder and cardiometabolic disease: improving causal inference to inform practice. *Psychological Medicine*, *47*(2), 209–225.

Koenen, Karestan C, Hitsman, B., Lyons, M. J., Stroud, L., Niaura, R., McCaffery, J., … Tsuang,

M. (2006). Posttraumatic stress disorder and late-onset smoking in the Vietnam era twin registry. *Journal of Consulting and Clinical Psychology*, *74*(1), 186–190.

Kraemer, H. C., Stice, E., Kazdin, A., Offord, D., & Kupfer, D. (2001). How do risk factors work together? Mediators, moderators, and independent, overlapping, and proxy risk factors. *The American Journal of Psychiatry*, *158*(6), 848–856.

Kral, V. A. (1951). Psychiatric observations under severe chronic stress. *The American Journal of Psychiatry*, *108*(3), 185–192.

Leblanc, M., & Crowley, J. (1993). Survival Trees by Goodness of Split. *Journal of the American Statistical Association*, *88*(422), 457–467.

Leor, J., Poole, W. K., & Kloner, R. A. (1996). Sudden cardiac death triggered by an earthquake. *N Engl J Med*, *334*(7), 413–419.

Li, K.-C. (1985). From Stein's Unbiased Risk Estimates to the Method of Generalized Cross Validation. *The Annals of Statistics*, *13*(4), 1352–1377.

Luque-Fernandez, M. A., Belot, A., Valeri, L., Cerulli, G., Maringe, C., & Rachet, B. (2018). Data-Adaptive Estimation for Double-Robust Methods in Population-Based Cancer Epidemiology: Risk Differences for Lung Cancer Mortality by Emergency Presentation. *American Journal of Epidemiology*.

MacKinnon, D. P. (2008). Introduction to statistical mediation analysis. *APA Handbook of Research Methods in Psychology Vol 2 Research Designs Quantitative Qualitative Neuropsychological and Biological*, 488.

MacKinnon, D. P., Kisbu-Sakarya, Y., & Gottschall, A. C. (2013). Developments in Mediation Analysis. In *The Oxford Handbook of Quantitative Methods in Psychology Volume 2 Statistical Analysis* (Vol. 2).

Madakasira, S., & O'Brien, K. F. (1987). Acute posttraumatic stress disorder in victims of a natural disaster. *The Journal of Nervous and Mental Disease*, *175*(5), 286–290.

McCauley, J. L., Killeen, T., Gros, D. F., Brady, K. T., & Back, S. E. (2012). Posttraumatic Stress Disorder and Co-Occurring Substance Use Disorders: Advances in Assessment and Treatment. *Clinical Psychology : A Publication of the Division of Clinical Psychology of the American Psychological Association*, *19*(3).

McFarlane, A. C. (1988). The phenomenology of posttraumatic stress disorders following a natural disaster. *The Journal of Nervous and Mental Disease*, *176*(1), 22–29.

Mogensen, U. B., Ishwaran, H., & Gerds, T. A. (2012). Evaluating Random Forests for Survival Analysis using Prediction Error Curves. *Journal of Statistical Software*, *50*(11), 1–23.

Morgan, L., Scourfield, J., Williams, D., Jasper, A., & Lewis, G. (2003). The Aberfan disaster: 33-Year follow-up of survivors. *British Journal of Psychiatry*, *182*(JUNE), 532–536.

Murphy, S. A. (1984). After Mount St. Helens. Disaster stress research. *Journal of Psychosocial Nursing and Mental Health Services*, *22*(7), 8–11, 15–18.

Naimi, A. I., & Balzer, L. B. (2018). Stacked generalization: an introduction to super learning. *European Journal of Epidemiology*.

Nakagawa, I., Nakamura, K., Oyama, M., Yamazaki, O., Ishigami, K., Tsuchiya, Y., & Yamamoto, M. (2009). Long-term effects of the Niigata-Chuetsu earthquake in Japan on acute myocardial infarction mortality: an analysis of death certificate data. *Heart (British Cardiac Society)*, *95*(24), 2009–2013.

Nakano, M., Kondo, M., Wakayama, Y., Kawana, A., Hasebe, Y., Shafee, M. A., … Shimokawa, H. (2012). Increased incidence of tachyarrhythmias and heart failure hospitalization in patients with implanted cardiac devices after the great East Japan earthquake disaster. *Circulation Journal : Official Journal of the Japanese Circulation Society*, *76*(5), 1283–1285.

Neria, Y., Galea, S., & Norris, F. H. (Eds.). (2009). *Mental Health and Disasters* (First). New York: Cambridge University Press.

Norris, F. H. (2006). Disaster research methods: Past progress and future directions. *Journal of*

Traumatic Stress, *19*(2), 173–184.

Norris, F. H., Friedman, M. J., Watson, P. J., Byrne, C. M., Diaz, E., & Kaniasty, K. (2002). 60,000 disaster victims speak: Part I. An empirical review of the empirical literature, 1981-2001. *Psychiatry*, *65*(3), 207–239.

Ogawa, K., Tsuji, I., Shiono, K., & Hisamichi, S. (2000). Increased acute myocardial infarction mortality following the 1995 Great Hanshin-Awaji earthquake in Japan. *International Journal of Epidemiology*, *29*(3), 449–455.

Pai, A., Suris, A. M., & North, C. S. (2017). Posttraumatic Stress Disorder in the DSM-5: Controversy, Change, and Conceptual Considerations. *Behavioral Sciences (Basel, Switzerland)*, *7*(1).

Pearl, J. (2001). *Direct and indirect effects. "Proceedings of the seventeenth conference on uncertainty in artificial intelligence"* (J. Breese & D. Koller, Eds.). San Francisco: Morgan Kaufmann Publishers Inc.

Pearl, Judea. (2010). On the consistency rule in causal inference: axiom, definition, assumption, or theorem? *Epidemiology (Cambridge, Mass.)*, *21*(6), 872–875.

Pearl, Judea. (2014a). Interpretation and identification of causal mediation. *Psychological Methods*, *19*(4), 459–481.

Pearl, Judea. (2014b). Is Scientific Knowledge Useful for Policy Analysis? A Peculiar Theorem Says: No. *Journal of Causal Inference*, *2*(1), 109–112.

Pirracchio, R., Petersen, M. L., Carone, M., Rigon, M. R., Chevret, S., & van der Laan, M. J. (2015). Mortality prediction in intensive care units with the Super ICU Learner Algorithm (SICULA): A population-based study. *The Lancet Respiratory Medicine*.

Polley, E.C., LeDell, E., & van der Laan, M. J. (2016). *SuperLearner: Super Learner Prediction (Version 2.0-23) [R Package].*

Polley, Eric C, & van der Laan, M. J. (2010). Super Learner in Prediction. *U.C. Berkeley Division*

of Biostatistics Working Paper.

Preacher, K. J., Rucker, D. D., & Hayes, A. F. (2007). Addressing moderated mediation hypotheses: theory, methods, and prescriptions. *Multivariate Behavioral Research*, *42*(1), 185–227.

Priebe, S., Marchi, F., Bini, L., Flego, M., Costa, A., & Galeazzi, G. (2011). Mental disorders, psychological symptoms and quality of life 8 years after an earthquake: Findings from a community sample in Italy. *Social Psychiatry and Psychiatric Epidemiology*, *46*(7), 615–621.

Ripoll Gallardo, A., Pacelli, B., Alesina, M., Serrone, D., Iacutone, G., Faggiano, F., … Allara, E. (2018). Medium- and long-term health effects of earthquakes in high-income countries: a systematic review and meta-analysis. *International Journal of Epidemiology*, *47*(4), 1317–1332.

Roberts, A. L., Agnew-Blais, J. C., Spiegelman, D., Kubzansky, L. D., Mason, S. M., Galea, S., … Koenen, K. C. (2015). Posttraumatic stress disorder and incidence of type 2 diabetes mellitus in a sample of women: a 22-year longitudinal study. *JAMA Psychiatry*, *72*(3), 203–210.

Robins, J. M. (1986). A new approach to causal inference in mortality studies with a sustained exposure period—application to control of the healthy worker survivor effect. *Mathematical Modelling*, *7*(9–12), 1393–1512.

Robins, J. M., & Greenland, S. (1992). Identifiability and exchangeability for direct and indirect effects. *Epidemiology (Cambridge, Mass.)*, *3*(2), 143–155.

Robins, J. M., Hernán, M. a, & Brumback, B. (2000). Marginal structural models and causal inference in epidemiology. *Epidemiology (Cambridge, Mass.)*, *11*(5), 550–560.

Rose, S. (2013). Mortality risk score prediction in an elderly population using machine learning. *American Journal of Epidemiology*.

Rubin, D. B. (2005). Causal Inference Using Potential Outcomes. *Journal of the American Statistical Association*, *100*(469), 322–331.

Salcioglu, E., Basoglu, M., & Livanou, M. (2003). Long-term psychological outcome for non-treatment-seeking earthquake survivors in Turkey. *The Journal of Nervous and Mental Disease*, *191*(3), 154–160.

Segal, M. R. (1988). Regression Trees for Censored Data. *Biometrics*, *44*(1), 35.

Singer, J. D., & Willett, J. B. (1993). It's About Time: Using Discrete-Time Survival Analysis to Study Duration and the Timing of Events. *Journal of Educational and Behavioral Statistics*.

Sobel, M. E. (1982). Asymptotic Confidence Intervals for Indirect Effects in Structural Equation Models. *Sociological Methodology*, *13*(1982), 290.

Sofia, S., Melone, A., Manzoli, L., De Ciantis, P., Varrato, E., Di Filippo, R., … Balsano, C. (2012). Cardiovascular and cerebrovascular events pre- and post-earthquake of 6 April 2009: the Abruzzo's experience. *American Journal of Hypertension*, *25*(5), 556–560.

Sokejima, S., Nakatani, Y., Kario, K., Kayaba, K., Minowa, M., & Kagamimori, S. (2004). Seismic intensity and risk of cerebrovascular stroke: 1995 Hanshin-Awaji earthquake. *Prehospital and Disaster Medicine*, *19*(4), 297–306.

Stone, M. (1977). An Asymptotic Equivalence of Choice of Model by Cross-Validation and Akaike's Criterion. *Journal Of The Royal Statistical Society Series B-Methodological*.

Tainaka, H., Oda, H., Nakamura, S., Tabuchi, T., Noda, T., & Mito, H. (1998). [Workers' stress after Hanshin-Awaji earthquake in 1995--symptoms related to stress after 18 months]. *Sangyō Eiseigaku Zasshi = Journal of Occupational Health*, *40*(6), 241–249.

Takegami, M., Miyamoto, Y., Yasuda, S., Nakai, M., Nishimura, K., Ogawa, H., … Naito, H. (2015). Comparison of Cardiovascular Mortality in the Great East Japan and the Great Hanshin-Awaji Earthquakes – A Large-Scale Data Analysis of Death Certificates –. *Circulation Journal*, *79*(5), 1000–1008.

Tsuchida, M., Kawashiri, M.-A., Teramoto, R., Takata, M., Sakata, K., Omi, W., … Yamagishi, M. (2009). Impact of severe earthquake on the occurrence of acute coronary syndrome and stroke in a rural area of Japan. *Circulation Journal : Official Journal of the Japanese Circulation Society*, *73*(7), 1243–1247.

U.S. Geological Survey. (2014). Earthquakes with 1,000 or More Deaths 1900-2014.

Valeri, L., & Vanderweele, T. J. (2013). Mediation analysis allowing for exposure-mediator interactions and causal interpretation: theoretical assumptions and implementation with SAS and SPSS macros. *Psychological Methods*, *18*(2), 137–150.

van den Berg, B, Grievink, L., van der Velden, P. G., Yzermans, C. J., Stellato, R. K., Lebret, E., & Brunekreef, B. (2008). Risk factors for physical symptoms after a disaster: a longitudinal study. *Psychological Medicine*, *38*(4), 499–510.

van den Berg, Bellis, Grievink, L., Yzermans, J., & Lebret, E. (2005). Medically unexplained physical symptoms in the aftermath of disasters. *Epidemiologic Reviews*, *27*, 92–106.

Van der Laan, M. J., & Dudoit, S. (2003). Unified cross-validation methodology for selection among estimators and a general cross-validated adaptive epsilon-net estimator: Finite sample oracle inequalities and examples. *U.C. Berkeley Division of Biostatistics Working Paper*.

Van der Laan, M. J., Dudoit, S., & Van der Vaart, A. W. (2006). The cross-validated adaptive epsilon-net estimator. *Statistics & Decisions*.

Van der Laan, M. J., Polley, E. C., & Hubbard, A. E. (2007a). Super learner. *Statistical Applications in Genetics and Molecular Biology*, *6*, Article25.

Van der Laan, M. J., Polley, E. C., & Hubbard, A. E. (2007b). Super learner. *Statistical Applications in Genetics and Molecular Biology*, *6*, Article25.

Vanderweele, T. J. (2015). Explanation in Causal Inference: Methods for Mediation and Interaction. In *Journal of Chemical Information and Modeling* (Vol. 53).

VanderWeele, T. J. (2014). A unification of mediation and interaction: a 4-way decomposition. *Epidemiology (Cambridge, Mass.)*, *25*(5), 749–761.

Vanderweele, T. J., Vansteelandt, S., & Robins, J. M. (2014). Effect decomposition in the presence of an exposure-induced mediator-outcome confounder. *Epidemiology (Cambridge, Mass.)*, *25*(2), 300–306.

Vanholder, R., van der Tol, A., De Smet, M., Hoste, E., Koç, M., Hussain, A., … Sever, M. S. (2007). Earthquakes and crush syndrome casualties: lessons learned from the Kashmir disaster. *Kidney International*, *71*(1), 17–23.

Vansteelandt, S., & Daniel, R. M. (2017). Interventional Effects for Mediation Analysis with Multiple Mediators. *Epidemiology (Cambridge, Mass.)*, *28*(2), 258–265.

Wang, A., & Arah, O. A. (2015). G-computation demonstration in causal mediation analysis. *European Journal of Epidemiology*, *30*(10), 1119–1127.

Watanabe, H., Kodama, M., Tanabe, N., Nakamura, Y., Nagai, T., Sato, M., … Aizawa, Y. (2008). Impact of earthquakes on risk for pulmonary embolism. *International Journal of Cardiology*, *129*(1), 152–154.

Watson, D. (2005). Rethinking the mood and anxiety disorders: a quantitative hierarchical model for DSM-V. *Journal of Abnormal Psychology*, *114*(4), 522–536.

Wessely, S., Nimnuan, C., & Sharpe, M. (1999). Functional somatic syndromes: One or many? *Lancet*, Vol. 354, pp. 936–939.

Wessely, Simon, & White, P. D. (2004). There is only one functional somatic syndrome. *British Journal of Psychiatry*, Vol. 185, pp. 95–96.

Wolpert, D. H. (1992). Stacked generalization. *Neural Networks*, *5*(2), 241–259.

Wright, S. (1923). The Theory of Path Coefficients a Reply to Niles's Criticism. *Genetics*, *8*(3), 239–255.

Wright, S. (1934). The Method of Path Coefficients. *The Annals of Mathematical Statistics*, *5*(3),

161–215.

Wu, T. Y., Cheung, J., Cole, D., & Fink, J. N. (2014). The Christchurch earthquake stroke incidence study. *Journal of Clinical Neuroscience : Official Journal of the Neurosurgical Society of Australasia*, *21*(3), 412–415.

Yashiro, M., Muso, E., Itoh-Ihara, T., Oyama, A., Hashimoto, K., Kawamura, T., … Sasayama, S. (2000). Significantly high regional morbidity of MPO-ANCA-related angitis and/or nephritis with respiratory tract involvement after the 1995 great earthquake in Kobe (Japan). *American Journal of Kidney Diseases : The Official Journal of the National Kidney Foundation*, *35*(5), 889–895.

Yzermans, C. J., van den Berg, B., & Dirkzwager, A. J. E. (2009). Physical Health Problems after Disasters. In Y. Neria, S. Galea, & F. H. Norris (Eds.), *Mental Health and Disasters* (pp. 67–93).

Zaetta, C., Santonastaso, P., & Favaro, A. (2011). Long-term physical and psychological effects of the Vajont disaster. *European Journal of Psychotraumatology*, *2*, 1–5.

Zhang, Z., Ran, M. S., Li, Y. H., Ou, G. J., Gong, R. R., Li, R. H., … Fang, D. Z. (2012). Prevalence of post-traumatic stress disorder among adolescents after the Wenchuan earthquake in China. *Psychological Medicine*.

Zheng, W., Balzer, L., van der Laan, M., & Petersen, M. (2018). Constrained binary classification using ensemble learning: an application to cost-efficient targeted PrEP strategies. *Statistics in Medicine*.

Zwiebach, L., Rhodes, J., & Roemer, L. (2010). *Resource Loss , Resource Gain , and Mental Health Among Survivors of Hurricane Katrina*. *23*(6), 751–758.